



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 :  C12Q 1/68, G01N 30/00	A1	(11) International Publication Number: WO 89/ 06703  (43) International Publication Date: 27 July 1989 (27.07.89)
---	----	---

(21) International Application Number: PCT/US89/00293  (22) International Filing Date: 23 January 1989 (23.01.89)  (31) Priority Application Number: 146,525  (32) Priority Date: 21 January 1988 (21.01.88)  (33) Priority Country: US  (71)(72) Applicants and Inventors: DRYJA, Thaddeus, P. [US/US]; 85 Forbes Road, Milton, MA 02143 (US). FRIEND, Stephen [US/US]; 14 Spencer Avenue, Somerville, MA 02143 (US). YANDELL, David, W. [US/US]; 53 Upton Road, Waltham, MA 02154 (US).  (74) Agent: CLARK, Paul, T.; Fish & Richardson, One Financial Center, Suite 2500, Boston, MA 02111-2658 (US).	(81) Designated States: AT (European patent), AU, BB, BE (European patent), BG, BR, CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, MW, NL (European patent), NO, RO, SD, SE (European patent), SU.
--	---

## Published

*With international search report.**Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

## (54) Title: DIAGNOSIS OF RETINOBLASTOMA

## (57) Abstract

This invention relates to nucleic acid, or fragments thereof, encoding the retinoblastoma polypeptide, the retinoblastoma polypeptide itself, methods of detecting a defective retinoblastoma gene in human patients, and methods of treating these patients.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT Austria	FR France	ML Mali
AU Australia	GA Gabon	MR Mauritania
BB Barbados	GB United Kingdom	MW Malawi
BE Belgium	HU Hungary	NL Netherlands
BG Bulgaria	IT Italy	NO Norway
BJ Benin	JP Japan	RO Romania
BR Brazil	KP Democratic People's Republic of Korea	SD Sudan
CF Central African Republic	KR Republic of Korea	SE Sweden
CG Congo	LI Liechtenstein	SN Senegal
CH Switzerland	LK Sri Lanka	SU Soviet Union
CM Cameroon	LU Luxembourg	TD Chad
DE Germany, Federal Republic of	MC Monaco	TG Togo
DK Denmark	MG Madagascar	US United States of America
FI Finland		

- 1 -

DIAGNOSIS OF RETINOBLASTOMA

Background of the Invention

This application is a continuation-in-part of Dryja et al. U.S.S.N. 146,525, filed January 21, 1988, which is a continuation-in-part of Dryja et al. U.S.S.N. 895,163, filed August 11, 1986, both hereby incorporated by reference.

This invention concerns the retinoblastoma gene and methods for detecting and treating patients afflicted with a defective retinoblastoma gene.

Retinoblastoma is a neoplastic condition of the retinal cells, observed almost exclusively in children between the ages of 0 and 4 years. It affects between 1 in 34,000 and 1 in 15,000 live births in the United States. (L.E. Zimmerman, 1985, Retinoblastoma and retinocytoma, In W.H. Spencer (ed.), Ophthalmic Pathology: an Atlas and Textbook, Vol. II, Philadelphia: W.B. Saunders Co., pp. 1292-1351.) If untreated, the malignant neoplastic retinal cells in the intraocular tumor travel to other parts of the body, forming foci of uncontrolled growth which are always fatal. The current treatment for a retinoblastoma is enucleation of the affected eye if the intraocular tumor is large; for small intraocular tumors, radiation therapy, laser therapy, or cryotherapy is preferred. There is no known successful treatment for metastatic retinoblastoma. As with most cancers, morbidity and mortality are reduced if diagnosis can be made early in the course of the disease.

- 2 -

In 30-40% of cases of retinoblastoma, the affected individual carries a heritable predisposition to retinoblastoma and can transmit this predisposition to his or her offspring as a dominant trait (A.G. Knudson, 1971, Mutation and cancer: Statistical study of retinoblastoma, Proc. Natl. Acad. Sci., Vol. 68, pp. 820-23). Carriers of this retinoblastoma-predisposing trait are at a greatly elevated risk for development of several other forms of primary cancer, notably osteosarcoma and soft-tissue sarcoma.

The genetic locus associated with familial retinoblastoma has been assigned to the q14 band of human chromosome 13 (R.S. Sparkes et al., 1980, Science, Vol. 208, pp. 1042-44). Most retinoblastomas arise from cells which have lost both normal, dominant, homologous alleles at this retinoblastoma locus. However, individuals carrying one defective allele may be predisposed to the disease. Children who have had one eye affected by retinoblastoma or who are related to someone with retinoblastoma may be genetically predisposed and therefore at risk of developing the disease. These individuals routinely are tested for retinoblastoma every 2-3 months by an ocular examination procedure which requires placing the child under general anesthesia.

#### Summary of the Invention

In general, the invention concerns purified nucleic acid (less than 100kb in size), and fragments thereof of at

- 3 -

least 15 bases, encoding the Rb gene. The invention also concerns cells transformed with this nucleic acid, isolated polypeptides encoded by this nucleic acid, and antibodies to this polypeptide, or to naturally occurring retinoblastoma polypeptide. Retinoblastoma polypeptide is the polypeptide encoded by the Rb gene. Further, the invention concerns a composition, suitable for treating a human having a defective Rb gene, containing retinoblastoma polypeptide, or a fragment thereof, in a pharmacologically acceptable carrier.

The invention also features methods of screening human patients to determine those not at risk of developing retinoblastoma and thus not requiring conventional examinations to be performed. This screening involves, for example, comparing nucleic acid of a patient with purified nucleic acid encoding a human Rb gene, or fragments thereof.

Thus in various aspects, the invention features methods of analyzing the predisposition of patients to retinoblastoma which involves detecting large and small deletions or point mutations in the retinoblastoma gene, or detecting the co-inheritance of such defects with specific restriction fragment length polymorphisms (RFLPs), or detecting the presence or absence of a normal or defective retinoblastoma gene by hybridizing a nucleic acid sample from the patient with a probe specific for the retinoblastoma gene, and determining the ability of the probe to hybridize to the nucleic acid. The

- 4 -

lack of hybridization to the nucleic acid indicates the presence of a large deletion in the gene. A probe specific for the retinoblastoma gene may be hybridized to fragments separated by a defined physical property from a sample of a patient, the hybrids of the probe and the fragments detected, and the hybrids compared to hybrids detected from the hybridization of the probe and separated nucleic acid fragments from a normal retinoblastoma gene. The absence of hybrids or presence of hybrids of a smaller size compared to a normal patient is an indication of large deletions in the retinoblastoma gene of the patient. Preferably, the probe specific for the retinoblastoma gene is the cloned DNA in p4.7R; or a fragment thereof; and the defined physical property is molecular weight.

Small deletions or point mutations can be detected by determining the nucleotide sequence of a retinoblastoma allele from a patient, and comparing the nucleotide sequence with the nucleotide sequence of a retinoblastoma allele, or subregion thereof, from a person not afflicted with retinoblastoma; or by detecting mismatches between a nucleic acid sample from a patient and a probe specific for the retinoblastoma gene from a person not afflicted with retinoblastoma. The co-inheritance of specific genetic polymorphisms with the retinoblastoma gene may be an indication of the predisposition of a patient to retinoblastoma. According to this method, nucleic acid

- 5 -

fragments are generated from a sample of the patient, the fragments are separated according to a defined physical property of the fragments (e.g., molecular weight), a detectable probe specific for the retinoblastoma gene is hybridized to the fragments, hybrids of the probe and the fragments are detected, and the hybrids are compared to hybrids detected from the hybridization of the same probe and separated nucleic acid fragments from a sample of a parent of the patient.

In another aspect, the invention features the use of an isolated normal human retinoblastoma gene to synthesize Rb polypeptide for use in the treatment of individuals determined to have a defective Rb allele.

In yet another aspect, the invention features a method of detecting the presence of the retinoblastoma polypeptide in a tumor sample from a human patient, by producing an antibody to the retinoblastoma polypeptide, contacting the antibody with the tumor sample, and detecting immune complexes as an indication of the presence in the tumor sample of the retinoblastoma polypeptide. The absence of the polypeptide indicates that the tumor is caused by a defect in a retinoblastoma allele. This procedure would preferably involve contacting a tumor sample from a human patient with an antibody (e.g., monoclonal antibody) which specifically reacts with the retinoblastoma polypeptide, or a fragment thereof, and determining whether the antibody binds to cells of the tissue

- 6 -

specimens. The absence of immune complexes is an indication that the tumor was the result of a defective retinoblastoma allele.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of the Preferred Embodiments

The description of the preferred embodiments of the invention will follow the brief description of the drawings given below.

Drawings

Fig. 1 is a diagrammatic representation of a restriction map of the insert in the clone p4.7R.

Fig. 2 is a diagrammatic representation of a restriction map of the genomic locus of the retinoblastoma gene.

Fig. 3 is a diagrammatic representation of the vectors p2AR3.8 and p2AR0.9 of the invention.

Fig. 4 is a scale map of the normal retinoblastoma gene.

Fig. 5 (5-1 through 5-3) is a nucleic acid sequence of a cDNA of the normal retinoblastoma gene, with flanking regions.

Fig. 6 (6-1 through 6-9) is a nucleic acid sequence of exons of the normal retinoblastoma gene, with flanking regions.

Fig. 7 is a restriction map of the retinoblastoma gene, showing the locations of DSPs.

- 7 -

Fig. 8 is a gel and a diagram showing the inheritance of the polymorphism RB1.3 in a retinoblastoma-prone family.

Fig. 9 is a diagram showing the segregation of the DSP RB1.3 in three families with hereditary retinoblastoma.

#### Retinoblastoma Polypeptide

The Rb polypeptide is the specific amino acid chain encoded by the nucleic acid sequence of the normal retinoblastoma gene. The Rb polypeptide of this invention includes: (1) naturally occurring retinoblastoma protein; (2) synthetically produced retinoblastoma polypeptide; and (3) retinoblastoma polypeptide produced from purified nucleic acid (e.g., cDNA or genomic DNA) via an in vitro expression system. Also included are biologically active fragments of Rb polypeptide which either have a biological activity of naturally occurring Rb polypeptide, or include an epitope of this polypeptide and thus are suitable for production of Rb-specific antibodies.

#### Retinoblastoma Gene

The Rb gene is that distinct nucleic acid sequence in the human genome, the absence or mutation of which predisposes one to retinoblastoma. The purified nucleic acid sequence encoding the retinoblastoma gene can be carried on vectors which can be propagated in cells. For the purposes of this invention, purified nucleic acid encoding the Rb gene is defined as nucleic acid isolated from its natural environment

- 8 -

(e.g., cDNA or a fragment of genomic DNA) which hybridizes specifically to the retinoblastoma gene under hybridizing conditions. An example of purified nucleic acid which encodes the retinoblastoma gene, and is carried on a vector, is the cDNA clone p4.7R. This clone was obtained in the following manner.

CDNA

The human DNA probe pH3-8, isolated from a human chromosome 13 lambda phage library (M. Lalande et al., 1984, Cancer Genet. Cytogenet., Vol. 13, pp. 283-95), was used in a chromosome walking technique to isolate and map 30 kilobases (kb) of genomic DNA surrounding the H3-8 sequence. One fragment generated by this technique, named p7H30.7R, was found to recognize a DNA sequence in the mouse genome as well as one within human chromosome 13 (T.P. Dryja et al., 1986, Proc. Natl. Acad. Sci. USA, Vol. 83, pp. 7391-94). The homology of p7H30.7R to both human and mouse DNA suggested that p7H30.7R contains coding sequences of a structural gene.

To test this possibility, p7H30.7R was radiolabeled and used to probe a Northern blot of RNA isolated from three retinoblastoma tumors and an adenovirus 12-transformed human embryonic retinal cell line (Vaessen et al., 1986, EMBO Journal, Vol. 5, pp. 335-). The p7H30.7R probe hybridized to an RNA transcript of approximately 4.7 kb from the retinal cell line, but did not hybridize to any RNA transcripts from the

- 9 -

three tumor samples.

Subsequently, RNA isolated from the adenovirus-transformed retinal cell line was used to construct a cDNA library. This library was screened with the labeled p7H30.7R probe. Several cDNA clones were isolated which had similar restriction maps. The longest of these, p4.7R, contained 4.7 kb of DNA. The restriction map of the insert in the clone p4.7R is shown in Fig. 1.

The p4.7R clone was used to screen RNA transcripts isolated from four retinoblastomas, an osteosarcoma, and the adenovirus-transformed retinal cells. In a Northern blot analysis of isolated RNA's, the p4.7R probe cross-reacted with a 4.7 kb transcript in the transformed retinal cells which was not present in the four retinoblastoma and one osteosarcoma cell samples.

#### Genomic DNA

Clones containing genomic DNA including the retinoblastoma gene were isolated in the manner described below. Recombinant bacteriophage libraries containing human genomic DNA fragments inserted in the lambda phage cloning vector EMBL-3 were constructed according to published methods (Seed et al., 1982, Gene, Vol. 19, pp. 201-209). Those recombinant bacteriophage which contain fragments of the retinoblastoma gene were initially detected by hybridization of the bacteriophage plaques with p4.7R.

- 10 -

Thirty-six distinct recombinant bacteriophage that contain overlapping human genomic DNA fragments were isolated. Selected bacteriophage were plaque-purified and amplified, and the restriction map of each phage insert was determined by the method of Rackwitz et al., Gene, Vol. 30, pp. 195-200.

With these bacteriophage a restriction map of a region that spans approximately 200 kb was constructed, shown in Figs. 2 and 4. All of the known sequences present in the mRNA from the retinoblastoma gene are present in this cloned region. In aggregate, the human DNA sequences in this set of bacteriophage represent the chromosomal segment of the Rb gene. In Fig. 2, the vertical marks above the map represent the location of HindIII sites, and the vertical marks below the map represent the location of EcoRI sites. The boxed areas represent HindIII fragments which contain sequences found in the cDNA (exons). Each double-headed arrow beneath the map represents a distinct recombinant bacteriophage clone. Fig. 4 shows all the recognition sites of the six restriction enzymes Hind III, EcoR I, Xba I, Sac I, Sac II, and BamH I (New England Biolabs, Inc.). Restriction endonuclease fragments that contained exons were identified by their hybridization with cDNA clones or synthesized oligonucleotide sequences based on cDNA sequence. These restriction fragments were subcloned in the plasmid vector Bluescribe (Stratagene, San Diego, CA). A total of 24 distinct plasmids were subcloned in this manner.

- 11 -

The number and size of each exon was determined by iterations of the following procedure. First, an oligonucleotide was synthesized that corresponded to the first 20 nucleotides of the cDNA sequence. Using this oligonucleotide as a primer, the plasmid with a genomic insert containing the most 5' exon was sequenced. The resultant sequence was aligned with the cDNA sequence to determine the length of the first exon; the point at which the plasmid and cDNA sequences diverged marked the beginning of the first intron. This exon and the flanking regions were further sequenced using synthetic oligonucleotide primers to generate a continuous nucleotide sequence composed of 5' promoter sequence, exon 1, and the beginning of intron 1. The second and subsequent exons were defined by synthesizing sequencing primers corresponding to the next 20 nucleotides of cDNA sequence that had not been previously assigned to an exon. All exons and the immediately adjacent flanking intron sequences were sequenced in both sense and antisense directions.

The dideoxynucleotide chain termination method of sequencing was carried out using the enzyme Sequenase (United States Biochemical Corporation, Cleveland, Ohio) according to protocols supplied by the manufacturer. The intron region downstream of exon 20 could not be sequenced by this method, due to an unusually problematic repeated sequence in this region that caused a series of 45 stops (bands appeared in all

- 12 -

four lanes of the sequencing gel). To resolve this region, sequencing reactions were carried out with Taq polymerase (Perkin-Elmer/Cetus). This enzyme allowed for the polymerization to be performed at 68°C and resolved the bases in this region.

All sequence data were analyzed and screened for overlapping regions using the sequence analysis program Microgenie Sequence Software (Beckman, Palo Alto, CA).

The position of each exon within the restriction map of the gene was determined by hybridization of cDNA fragments or synthetic oligomer sequences to recombinant bacteriophage DNA that had been digested with various restriction endonucleases. The precise location of most exons was subsequently deduced when recognition sequences of endonucleases were identified within the intron-exon sequence and correlated with the map. The position of each of the remaining exons was arbitrarily placed in the center of the smallest restriction fragment to which it hybridized.

The organization of the 27 exons along the genomic map of the retinoblastoma gene is illustrated in figure 4. This figure details the recognition sites for 6 restriction endonucleases and the position of the exons relative to these sites. Exons 1, 2, 3, 6, 9, 10, 13, 21, 22, 23, 24, 235, 26, and 27 have been precisely localized on this map. The other exons were mapped within small restriction fragments and are

- 13 -

illustrated in the middle of these fragments. Exons 11, 17, and the cluster of exons 14-16 were mapped by this technique with uncertainties of not more than 2.0 kb. The remaining exons (exons 4, 5, 7, 8, 12, 18, 19, and 20) were all mapped to within 0.8 kb of their true locations. For reference, this map also shows the position of several naturally occurring restriction fragment length polymorphisms.

Fig. 6 (6-1 through 6-9) shows the sequence flanking and including each exon. The exons range in size from 31 nucleotides (exon 24), to 1973 nucleotides (exon 27). The shortest intron sequence was found to be 80 nucleotides long and is located between exons 15 and 16, whereas the largest spans approximately 70.5 kb between exon 17 and 18. All of the intron donor and acceptor splicing sites comply with the GT-AG splice junction rule. Our methods of sequencing proved more accurate than previous reports in defining the exact number of exons comprising the retinoblastoma gene.

The first exon and the region immediately 5' to this exon are very G-C rich, which is a characteristic of promoter regions. This region contains 9 possible Hpa II restriction sites, is composed of 66% C+G nucleotides and does not exhibit CpG suppression. These criteria are indicative of a HTF island. This promoter region contained some nucleotides that could not be resolved in either the sense or antisense direction using either Sequenase or Taq polymerase.

- 14 -

Presumably, this was due to secondary structure that forms in this promoter region.

Analysis of the sequence approximately 30 nucleotides upstream of the transcription initiation site defined by Lee et al. (1987b) does not reveal a TATA box that is found in other promoter regions. This suggests that either the previously published initiation site is not in fact correctly defined, or that the retinoblastoma gene lacks the prototypical TATA and CAAT boxes of promoter regions. Further analysis of the sequence 5' to exon 1 reveals a possible TATA box at base pair #-274, labeled base pair #122 in figure 6-1. Homology for the seven base region is only 57%, yet the first four bases T-A-T-A, which are the bases most frequently conserved, are 100% homologous. A possible capping site, CAC, is located 14 bases away and again 49 bases away. CAAT boxes were not identified although the region is generally G-C rich.

The intron sequence that flanks the 3' side of exon 20 consisted of 21 consecutive repeats of the sequence TTT(T)C that together span 87 nucleotides. The number of repeat units can vary between different individuals and the alleles determined by the number of repeats behave like a heritable DNA polymorphism.

A computer search of the sequence data identified several intron regions homologous to Alu repetitive sequences. Alu repeats were located in the following regions: (1)

- 15 -

downstream of exon 2, between bp 492 and bp 704 according to the numbering scheme in Figure 6; 2) upstream of exon 9, between bp 19 and bp 117; 3) downstream of exon 11, between bp 504 and bp 680; 4) the intron sequence flanking both sides of exon 14 between bp 132 and bp 270 and between bp 420 and bp 741; and, 5) upstream of exon 17, between bp 36 to bp 210. The Alu sequence located downstream of exon 2 contains two internal sequences that are highly conserved in Alu repetitive sequences. The first is a sequence (GAGGCNGAGC) corresponding to the T-antigen binding sequence of the SV40 replication origin. The second is a symmetrical sequence (CCAGCCTGG) of no known function. This short symmetrical sequence is also present in both of the Alu sequences on either side of exon 14. Exons 14 and 15 are separated by a short intron that is almost entirely composed of Alu sequence, suggesting that exons 14 and 15 were possibly at one time a single exon and were divided by the insertion of an Alu element during evolution. This Alu sequence may have been directed to this position by the other Alu sequence located on the 5' end of exon 14 because retroposons have a tendency to integrate adjacent to one another.

The 3' end of the retinoblastoma gene contains the usual polyadenylation signal sequence, AATAAA. One sequence (TGTGTTCT) located 32 bases downstream of this hexamer is equivalent to the conserved downstream consensus sequence

- 16 -

(YGTGTYY) described by McLauchlan et al. (1985). This sequence and surrounding bases compose the "G/T cluster" generally found in a region 30 nucleotides downstream of the polyadenylation signal sequence (Birnstiel et al., 1985).

The p4.7R probe also was used to screen genomic DNA isolated from the tumors of 50 unrelated individuals (40 retinoblastomas, 8 osteosarcomas, and 2 undifferentiated tumors of unknown cellular origin arising in patients with hereditary retinoblastoma), as described in more detail below. These DNA samples were digested with HindIII and analyzed by Southern blot hybridization using radiolabeled p4.7R as the probe. This analysis revealed three types of deviant patterns of the genomic DNA restriction fragments: totally absent fragments, representing apparent homozygous deletions; under-represented fragments, representing apparent heterozygous deletions; and fragments of altered size, reflecting either partial deletion or an alteration of a restriction site. At least 30% of the tumor DNA's exhibited one of these abnormalities. In comparison, Southern blot analysis of leucocyte DNA from 18 normal individuals showed a uniform pattern of restriction fragments.

#### Use

The cDNA and genomic sequences, e.g., those in p4.7R, can be used, according to the invention, to screen individuals for the presence of a mutated allele of the Rb gene. This

- 17 -

screening procedure will allow individuals having a risk of developing retinoblastoma--because of family history or a previous incidence of retinoblastoma in one eye--to determine the need for routine testing by the current ocular examination procedure. Only if the screening procedure determines that the individual possesses a mutant Rb allele will the examination procedure need to be conducted on a regular basis. Those with two normal Rb alleles can discontinue examination, as the risk of developing retinoblastoma in an individual with two normal copies of the Rb gene is approximately 1 in 20,000, or 0.005%, compared to a risk of 80%-90% if an individual has an Rb allele containing a mutation sufficient to inactivate the allele. Thus, a substantial percentage of individuals who are currently examined regularly are not actually at a greater risk than the general population: neither a family history of nor a previous incidence of retinoblastoma is conclusive evidence that an individual has the genetic predisposition to the disease. Therefore, such individuals, actually carrying two normal copies of the Rb gene, have been repeatedly undergoing the expensive and traumatic ocular examination procedure needlessly.

The screening procedure according to the invention includes: (1) testing a nucleic acid sample of a patient for large deletions in the Rb gene locus; (2) testing a nucleic acid sample of a patient for small deletions or point mutations in the Rb gene locus; and (3) testing a nucleic acid sample of

- 18 -

a patient for RFLPs linked to the Rb gene locus.

Detection of Large Deletions in the Rb Gene

The availability of DNA probes from the Rb gene provides a means of directly detecting genetic lesions that create retinoblastoma-predisposing alleles. Suitable probes include the entire normal retinoblastoma gene sequence, or fragments thereof consisting of 15 or more bases encoding a specific portion of the retinoblastoma gene. When performed by Southern blot and dot blot procedures, this analysis is generally limited to the study of those lesions that create gross structural changes in the Rb gene, such as deletion of many hundreds of base pairs.

The DNA for a Southern Blot or dot blot analysis is isolated from peripheral leucocytes or, if the patient has had a tumor in one eye, from the tumor. To examine leucocyte DNA, a 10 ml blood sample is obtained from the individual, and the genomic DNA is isolated from the leucocytes in the sample, according to standard techniques. This DNA is digested with a restriction endonuclease (e.g., HindIII), and the resulting fragments are separated on an agarose electrophoresis gel according to a physical property such as molecular shape or molecular weight. For the purposes of this invention, molecular shape is defined as the structural configuration of the molecule (e.g., linear, circular, double-stranded or single-stranded). The DNA in the gel is transferred to a

- 19 -

nitrocellulose filter by blotting. The filter is then probed with, e.g., radiolabeled p2AR3.8 and, separately, p2AR0.9, containing subfragments from p4.7R obtained by EcoR1 digestion. (The diagrams of the vectors p2AR3.8 and p2AR0.9 are shown in Figure 3.) In order to more precisely define the location of any abnormalities detected, two or more subfragment probes are used separately rather than the entire p4.7R insert probe. The autoradiograms of the probed filter generate the data necessary to construct a restriction map of the Rb locus in the somatic or tumor DNA of the tested individual.

This restriction map is compared with a control restriction map, determined by using the same restriction enzymes for digestion and the same probe. A suitable control is DNA obtained from an adenovirus-transformed retinal cell line or leucocyte DNA from a set of normal individuals. If the tested individual has an Rb allele containing a significantly large deletion, a restriction map of his DNA, compared with the control, will contain an additional band or bands, and/or a band or bands that have lost 50% of their intensity, caused by a change in the size, or total elimination, of one or more restriction fragments by the deletion in one allele at the Rb locus.

This screening procedure by Southern analysis will detect the existence of Rb alleles which have large deletions and are thereby non-functional. If this analysis indicates

- 20 -

that the tested DNA from an individual has a restriction map which is different from the control map, there is a high probability that the individual contains a non-functional, mutant Rb allele. The individual must be monitored closely thereafter for the development of retinoblastoma.

If the test restriction map appears identical to the control, a different screening procedure can be performed to determine if the individual possesses an Rb allele having a small deletion or point mutation. Small deletions and point mutations may be sufficient to inactivate the allele, but not prevent hybridization with a probe. An example of this screening procedure is outlined below.

Detection of Other Mutations in the Rb Gene

To examine a DNA sample of an individual for small deletions or point mutations in the Rb locus, both homologs of the Rb gene from said individual are cloned. The cloned alleles then can be tested for the presence of nucleic acid sequence differences from the normal allele, e.g., as represented by p4.7R, by one of the following two methods: (1) the nucleotide sequence of both the cloned alleles and p4.7R are determined and then compared, or (2) the RNA transcripts from p4.7R are hybridized to single stranded whole genomic DNA from an individual to be tested, and the resulting heteroduplex is treated with Ribonuclease A (RNase A) and run on a denaturing gel to detect the location of any mismatches. In

- 21 -

more detail, these methods can be carried out according to the following procedure.

The alleles of the Rb gene in an individual to be tested are cloned using conventional techniques. A common method, for example, employs the bacteriophage vector EMBL3 (Frischauf et al., 1983, J. Mol. Biol., Vol. 170, pp. 827-). A 10 ml blood sample is obtained from the individual. The genomic DNA isolated from the cells in this sample is partially digested with MboI to an average fragment size of approximately 20 kb. Fragments in the range from 18-21 kb are isolated. The resulting MboI-ended fragments are ligated into the EMBL3 vector DNA which has been completely digested with BamHI, treated with alkaline phosphatase, and heated to 68°C for 10 minutes to disrupt the cohesive ends. This ligation mix is used in an in vitro lambda packaging reaction, and the packaged phage are amplified by growing a plate stock. (This cloning technique is described generally in Maniatis et al., 1982, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Publications, pp. 256-293.)

Approximately  $5 \times 10^5$  plaque forming units (pfu) from this plate stock are then screened with radiolabeled p4.7R by hybridization and autoradiography. Plaques which show hybridization to the p4.7R probe are plaque-purified and rescreened according to the above procedure. Positive plaques from the rescreening are isolated and used to prepare DNA

- 22 -

putatively containing Rb alleles from the individual.

The MboI genomic inserts in these isolated EMBL3 vector DNA samples are tested for the location of the sequences homologous to p4.7R by Southern analysis. DNA samples containing the entire Rb gene region are selected, and the appropriate restriction fragments containing the Rb gene from these samples are subcloned into a suitable vector, such as pUC9. These subclones thus contain copies of one or both Rb alleles from the DNA of the individual to be tested. To determine if both alleles are represented, the initial phage isolates are tested for the existence of restriction polymorphisms. These subcloned alleles are then examined for differences from p4.7R by one of the following techniques.

First, the nucleotide sequence of the normal Rb gene in p4.7R is determined. Restriction fragments of approximately 500 base pairs (bp) from p4.7R are subcloned into an M13mp8 phage vector and sequenced by the dideoxy technique (Sanger et al., 1977, Proc. Natl. Acad. Sci. USA, Vol. 74, pp. 5463-). A composite sequence of the Rb gene then can be assembled from these individual subclone sequences. The complete sequence of the normal retinoblastoma gene and flanking sequences is shown in Fig. 5 and Fig. 6.

The isolated Rb gene alleles are sequenced according to the following procedure. Restriction fragments (~ 2kb) of the allele are subcloned into the M13mp8 vector, and short

- 23 -

stretches (~500 bp) are sequenced individually using small restriction fragments isolated from p4.7R as the primers in the dideoxy sequencing reactions. The composite nucleotide sequence of the isolated allele then can be constructed from these individually-primed sequences. This sequence is compared directly with the sequence of the normal Rb gene, determined from p4.7R, to reveal any deletions or point mutations in the isolated allele.

An alternative method of comparing the allelic DNA with the normal Rb gene employs RNase A to assist in the detection of differences between the p4.7R sequence and the allele sequence. This comparison is performed in steps using small (~500 bp) restriction fragments of p4.7R as the probe. First, p4.7R is digested with a restriction enzyme(s) that cuts the Rb gene sequence into fragments of approximately 500bp. These fragement are separated on an electrophoresis gel, purified from the gel and cloned individually, in both orientations, into an SP6 vector (e.g., pSP64 or pSP65; Melton et al., 1984, Nucleic Acids Res., Vol. 12, pp. 7035-). The SP6-based plasmids containing inserts of p4.7R fragments are transcribed in vitro using the SP6 transcription system, well known in the art, in the presence of [ $\alpha$ -<sup>32</sup>P]GTP, generating radiolabeled RNA transcripts of both strands of the cDNA of the Rb gene.

Individually, these RNA transcripts are used to form

- 24 -

heteroduplexes with the allelic DNA, as described by Myers et al., 1985, Science, Vol. 230, pp. 1242-46, the teachings of which are incorporated herein by reference. Mismatches that occur in the RNA:DNA heteroduplex, owing to sequence differences between the p4.7R fragment and the Rb allele subclone from the individual, result in cleavage in the RNA strand when treated with RNase A. Such mismatches can be the result of point mutations or small deletions in the individual's Rb allele. Cleavage of the RNA strand yields two or more small RNA fragments, which run faster on the denaturing gel than the RNA probe itself.

In the RNase A technique, radiolabeled Rb gene RNA is hybridized to single strands of an individual's Rb alleles which have been cloned into a vector. The RNase A technique is advantageous, however, because it also can be used without having to clone the Rb alleles. Preferably, genomic DNA is isolated from blood cells of the individual to be tested, and this genomic DNA is hybridized directly with the radiolabeled Rb RNA probes to determine sequence differences from the normal Rb gene. Specifically, 5 µg of isolated, total genomic DNA is resuspended with the labeled RNA probe in 30 µl of hybridization buffer (80% formamide, 40mM Pipes pH6.4, 0.4M NaCl, and 1mM EDTA), and this hybridization mix is treated at 90°C for 10 minutes to denature the DNA. The mixture then is cooled slowly to 45°C and incubated at this temperature for 10

- 25 -

hours to allow hybridization of the RNA probe to the single-stranded DNA copies of the Rb allele. After hybridization, RNase A treatment and electrophoresis are performed as described by Myers et al., supra. Mismatches between the RNA probe and the genomic copies of the individual's Rb alleles are then readily detected.

Detection of RFLPs Linked to the Rb Gene

The inheritance of a retinoblastoma- predisposing defect can be traced by following its co-inheritance with DNA polymorphisms in a pedigree analysis.

The gene map shown in Figure 2 was used to develop nucleic acid probes useful for retinoblastoma diagnosis. To do so, the bacteriophage DNA corresponding to the human inserts were subcloned in the plasmid vector "Bluescribe" (Stratagene). Fifteen single-copy DNA fragments from the gene, ranging in size from 500 bp to 2000 bp, were subcloned. These sequences are scattered over the 200 kb of the mapped region. Subcloned DNA fragments were separated from vector sequences by digestion of plasmid DNA with one or more restriction endonucleases, electrophoresis through a 0.6% low-melting-point agarose gel, and purification by chromatography using an Elutip-d column (Schleicher and Schuell). Purified DNA fragments were radiolabeled with  $^{32}\text{P}$ -dCTP (New England Nuclear) by the random primer technique using the Klenow fragment of DNA polymerase I.

- 26 -

Restriction fragment length polymorphisms (RFLP's) were discovered by digesting genomic DNA isolated from six normal individuals with 33 different restriction enzymes. The DNA fragments resulting from the 198 separate digests were separated on a 0.8% agarose electrophoresis gel according to molecular shape or molecular weight. The DNA was transferred to nitrocellulose filters and hybridized with single copy DNA probes purified from the retinoblastoma gene according to published methods (T.P. Dryja et al., 1986, Proc. Natl. Acad. Sci. USA, Vol. 83, pp. 7391-94).

Of the 15 single-copy probe fragments, only five reveal RFLP's. Four of the polymorphisms appear to be the result of minor alterations (perhaps single base changes) in the recognition sequence of a restriction endonuclease (KpnI, XbaI, MboII, or Tth111I). The fifth polymorphism reflects variability in the number of tandem repeats of a 50 base pair sequence. The location of the DNA polymorphisms are shown in the map in Figure 2 (vertical arrows above the map). The location of the polymorphic MboII site(s) has not been determined precisely but is located at approximately 175 kb on this map. The frequencies of alleles which correspond to particular DNA polymorphisms are indicated in Table 1.

In order to demonstrate the utility of these probes to detect the presence of retinoblastoma-predisposing alleles in humans, twenty pedigrees with hereditary retinoblastoma were

- 27 -

analyzed. DNA was extracted from leucocyte nuclei of venous blood from available family members according to the method of Kunkel et al., 1977, Proc. Natl. Acad. Sci. USA, Vol. 74, pp. 1245-49, hereby incorporated by reference. For analysis of a kindred with a given RFLP, DNA from the available family members was digested with the appropriate restriction endonuclease. The resulting fragments were separated by agarose-gel electrophoresis, transferred to nitrocellulose filters, and hybridized to labeled probe.

In these families, the inheritance of alleles determined by the DNA polymorphisms within the retinoblastoma gene were traced and compared to the inheritance of the retinoblastoma-predisposing trait. For example, consider the polymorphism detected by probe p68RS2.0 (see Table 1). When genomic DNA is digested with the restriction enzyme RsaI, this probe hybridizes to allelic DNA fragments of different lengths. The size of these fragments ranges from 1.5 kb to 2.0 kb with intervals of approximately 50 bp. The DNA sequence of the 2.0 kb genomic fragment cloned in p68RS2.0 has a 50 to 53 bp segment which is repeated approximately 30 times (Table 2). This 53bp segment can be used as a probe in these analyses. A portion of the repeated sequence has homology to core sequences of VNTR's (Variable Number of Tandem Repeat) reported elsewhere (Y. Nakamura et al., 1987, Science, Vol. 235, pp. 1616-22). (The 11 bp sequence shown in Table 2 above the repeat unit



- 29 -

TABLE 1  
DNA POLYMORPHISMS

DNA PROBE	RESTRICTION ENDONUCLEASE	IDENTIFIED WITHIN THE RETINOBLASTOMA GENE LOCATION <sup>1</sup>	MAP SIZE (kb) <sup>3</sup>	ALLELES FREQUENCY <sup>4</sup>
p68RS2.0	<u>Rsa</u> I	142-143 kb	2.00 1.95 1.90 1.85 1.80 1.75 1.65 1.50	0.13 0.02 0.07 0.07 0.35 0.20 0.09 0.09
p88PR0.6 <sup>5</sup>	<u>Xba</u> I	120 kb	7.0 5.5	0.55 0.45
p35R0.6	<u>Tth</u> 111I	195 kb	4.95 4.35	0.20 0.80
p2P0.3	<u>Mbo</u> II	175 kb <sup>2</sup>	1.0 0.8 0.6 0.3	>.90 <.05 <.05 <.05
p95HS0.5	<u>Kpn</u> I	25 kb	12.0 8.0	0.95 0.05

<sup>1</sup>The map location of each polymorphic site refers to the position on the restriction map of the gene shown in Figure 1.

<sup>2</sup>The location of the MboII site is approximate, since the precise position of this site within the map is not yet known.

<sup>3</sup>The allele sizes were calculated from several independent measurements using HindIII fragments of lambda phage DNA as a standard.

<sup>4</sup>Allele frequencies are based on a population of 40-60 unrelated individuals.

<sup>5</sup>p88PRO.6, p35R0.6, p2P0.3 and p95HS0.5 are probes isolated from other regions of the retinoblastoma gene, shown in Fig. 2.

- 30 -

represents the core sequence reported for some VNTR's observed by Nakamura et al.) Because such tandemly repeated sequences tend to be genetically unstable, the number of repeats is highly variable. Eight distinct alleles at this site have been detected, and more may exist. Because of this number of common alleles, seventy-five percent of unrelated individuals are heterozygous for this polymorphism. The high frequency of heterozygosity makes this polymorphism extremely useful. (In Table 2, the brackets "(" and ")" denote regions of variability within the repeat unit; and the bases underlined above and below the bracketed regions denote possible alternate bases for those regions of variability.)

TABLE 2  
SEQUENCE OF THE REPEAT UNIT WITHIN p68RS2.0.

TG	A	GGGNNGTGGGG
CGTGTTGAA(	<u>A</u>	
<u>G</u>	ACAC( <u>C</u> )	AG
	TCCCCAATGCTGGAGGTGAGGTTGGT( <u>AGAAG</u> )ATGACT	

Fourteen retinoblastoma families carried constellations of alleles at this DNA polymorphism. This variation allows an examination of the frequency of co-inheritance of this site with the retinoblastoma-predisposing trait. For example, genomic DNA of individuals in these families was digested with Rsa I and the co-inheritance of any one Rsa I fragment with defective Rb alleles determined.

- 31 -

### Sequence Analysis

RFLP analysis can reveal only those sequence variations that give rise to a restriction fragment that is detectably different on a Southern blot from restriction fragments characterizing the normal nucleic acid. Most such detectable polymorphisms result from DNA sequence variation within a restriction endonuclease recognition site. These sites are rare, and finding them requires a laborious and expensive screening process in which genomic DNAs from several unrelated individuals are digested with as many as 50 different restriction enzymes. The fraction of all genomic DNA sequence polymorphisms (DSPs) at a specific locus that can be detected as RFLPs is small and depends on the number of enzymes used for screening; generally 90% or more of the DNA sequence polymorphism in the human genome is not within reach of RFLP-based analysis. Tens or even hundreds of potentially useful DSPs may exist within or near most disease-causing genes, but often only a few and sometimes none of these DSPs are detectable as RFLPs. Fig. 4 shows a scale map of the 200 kilobase genomic region that includes the 27 exons of the human retinoblastoma gene. The 27 exons make up a 4.7 kilobase transcript.

Fig. 7 shows the locations of the DNA sequence polymorphisms (DSPs) identified in this gene. Polymorphisms identified by the name of a restriction enzyme are RFLPs;

- 32 -

polymorphisms RB1.2, RB1.3, RB1.20 and RB1.26 are not detectable as RFLPs, and were found by PCR-amplification and direct sequencing, as described below. The 200 kilobase genomic region was isolated in a series of overlapping inserts from 35 distinct recombinant bacteriophage lambda clones. Exon-containing segments were subcloned into bluescribe plasmid cloning vectors (Stratagene, Inc.). Initial sequencing of cloned plasmid inserts was carried out using conventional methods for plasmid sequencing. Based on this genomic sequence, pairs of 20-base oligonucleotide primers were synthesized so that numerous regions 320 - 1200 bp in size could be amplified from genomic DNA by the polymerase chain reaction (PCR) method of Mullis et al.

For each amplification reaction, from 200 ng to 1.0 ug of genomic DNA was prepared in a reaction buffer containing 20 mM Tris (pH 8.4 or pH 8.6), 30 ug/ml bovine serum albumin, 300 mM to 7.5 mM, 10-50 pM of each oligonucleotide primer, and 1 unit of Taq polymerase (Perkin-Elmer Cetus). Optimal MgCl<sub>2</sub> concentrations and pH of the PCR reactions varied depending on the primer pair. PCR-amplification (30-35 rounds) was carried out following a cycle of 10 seconds at 94°C (denaturation), 10 seconds at 42-55°C (annealing), and 30 seconds at 70°C (Polymerization) using a programmable thermal cycler (Ericomp Corp., San Diego). All times are based on sample temperature rather than heat-block temperature, and do not include 'ramping

- 33 -

time' for the heat block. Optimal annealing temperatures varied for each primer pair. This protocol allowed PCR-amplification of regions as large as 2500 bp. Genomic DNA from 9-20 individuals was amplified for each region and screened for DSP.

In order to detect and utilize a greater fraction of the existing DNA sequence polymorphism in the retinoblastoma gene, we applied techniques of polymerase chain reaction (PCR), generally as described by K.B. Mullis et al., 1987, Methods Enzymol., Vol. 155, pp. 335-51, and direct sequencing, generally as described by C. Wong et al., 1987, Nature, Vol. 330, 384-86, to analyse normal allelic variation at this locus. Oligonucleotide primers were synthesized to amplify regions from the gene that varied in size from 320 - 1200 bp. Amplification and sequencing were carried out on DNA from at least 9 unrelated individuals for all regions screened, though for many regions 15 or more individuals were analyzed. In most cases, primer pairs were derived from intronic sequences that flanked one of the 27 exons of the gene, such that the PCR-amplified region contained both intron and exon sequences.

The results of this screening process are shown in Tables 3 and 4. Amplified DNA sequences were compared to one another and checked against sequence data from previously cloned plasmid inserts derived from the same region. Bases obscured by technical artifacts or other ambiguities were not

- 34 -

tabulated. Of 3712 bp of genomic DNA sequence screened at this locus, four sequence variations were identified (Table 3; map locations are shown in Fig. 7). All four variations were found in introns; of these, one is likely a rare variant (found in only 1 or 15 individuals sequenced), and three represent bona fide DNA sequence polymorphisms. A representative example (RB1.3) is illustrated in Fig. 7. This polymorphism occurs near exon 3 of the retinoblastoma gene. Neither form of the polymorphic sequence forms the recognition site of a known restriction enzyme, and hence this DSP is not detectable as an RFLP. Among a total of 82 genetically distinct (from unrelated individuals) alleles examined, no other base was observed at this site.

Fig. 8 illustrates the inheritance of the polymorphism RB1.3 in a retinoblastoma-prone family. Oligonucleotide primers (see Table 4) were used to PCR-amplify a 530 bp region of the human retinoblastoma gene that includes exon 3. The amplified fragment was sequenced by the methods described below. The sequence surrounding the polymorphism is written at the left side of the figure, read 5' to 3' from bottom to top, and the polymorphic bases are identified by adjacent tic marks.

The details of the analysis were as follows. Prior to sequencing, all PCR-amplified DNA samples were treated with proteinase-K and extracted with phenol/chloroform. High molecular weight DNA was separated from unused dNTPs and

- 35 -

oligonucleotide primers by column purification through sepharose CL-6B (Pharmacia). 250-400 ng of double-stranded PCR-amplified template was combined with 1-2 pM of (<sup>32</sup>P) end-labeled sequencing primer, and heat-denatured for 3 minutes at 96°C. This primer-template mixture was added to a buffer containing: MgCL<sub>2</sub> (2.5 mM), Tris-HCL pH 7.5 (5 mM), 6 units Sequenase (U.S. Biochemical) and dithiothreitol (3 mM), and divided into 4 reaction mixtures each containing all four deoxynucleotides (32 uM each) and one dideoxynucleotide (5 uM). This mixture was immediately incubated for 5 minutes at 37-42°C, and polymerization was stopped with a 0.37% EDTA stop buffer. Prior to loading on sequencing gels, the samples were heat denatured at 96°C for 2 minutes. Conventional 0.4 mm thick, 6% polyacrylamide sequencing gels were used, and autoradiography was typically for 12-24 hours without an intensifying screen.

Fig. 9 illustrates segregation of the DSP RB1.3 in three families with hereditary retinoblastoma. Alleles are shown beneath the symbol for each person. Affected individuals are indicated by filled symbols. In family RB-32, the (-) allele is the result of an intragenic deletion. By subsequent Southern blotting studies, the deletion was found to extend from exon 2 to exon 17. Based on these results, it can be predicted that the unaffected members of family RB-32 who carry the (G,-) genotype are also carriers of the mutation.

- 36 -

The DSPs we have detected are valuable genetic markers for our studies of hereditary retinoblastoma. In its hereditary form, a predisposition to the disease is passed from affected individuals to their offspring as a dominant trait with 90% penetrance. It can be seen from Fig. 8 that the affected father, who has passed the disease to two children, is heterozygous for RB1.3. Both affected children received the G allele, while the unaffected child inherited the allele marked by an A at this polymorphic site. In this family, then, inheritance of the G allele from the affected parent is in phase with and diagnostic for the disease-predisposing phenotype. Fig. 9 shows our analysis of three other retinoblastoma-prone families using RB1.3. Inheritance of the polymorphic markers we describe here has followed the expected Mendelian pattern in every family examined so far. No cross-overs were observed between the polymorphic sites and the retinoblastoma-predisposing trait in any of the pedigrees. This follows our expectations since the polymorphisms are within the disease gene. In family RB-32, an intragenic deletion in one copy of the RB gene, presumably causing the predisposition to the tumor, was identified by Southern blotting (data not shown). The deletion includes the region surrounding RB1.3, and hence carriers of the disease-predisposing allele are genotypically hemizygous for the A allele (A,-). Two unaffected members of pedigree RB-32

- 37 -

are carriers for the disease-predisposition, based on analysis of RB1.3 (see Fig. 9). More happily, Fig. 9 shows that the other unaffected children in pedigree RB-32, as well as those in pedigrees RB-36 and RB-50, are not carriers of the cancer-predisposition and therefore will not pass the disease on to their children. These results highlight the diagnostic value of this class of human genetic markers that were heretofore unavailable for this purpose.

The data we present may also be used to estimate the level of heterozygosity in the human genome from a novel perspective. Previous estimates based on restriction enzyme screening may be subject to a bias because the sequences recognized by these enzymes do not necessarily reflect a random sampling. It is likely that a substantially higher level of polymorphism occurs at CpG pairs than elsewhere. This is reflected by the relatively high proportion of RFLPs revealed by such enzymes as Msp I (CCGG) and Taq I (TCGA). The method we describe is not subject to this bias. From the results of our screening, it can be calculated that genomic heterozygosity at this locus is approximately  $h=0.00039$ . If only the intron sequences are considered, this estimate increases to  $h=0.00070$ . These estimates are below the predictions of others, and may reflect the absence from our methods of the bias described above. However, an analogous calculation of heterozygosity ( $0.00044 < h < 0.00087$ ) based instead on our

- 38 -

initial RFLP screening is also below the estimates of others and is quite consistent with our estimate based on direct sequencing. It seems likely that the human retinoblastoma gene is intrinsically less polymorphic than many other regions of the genome. Although mutations in this gene are known to be early events in the formation of several types of cancer, it is unclear why polymorphism at this locus may have been selected against in human evolution.

The approach for detecting DSPs demonstrated here has several advantages over conventional RFLP-based screening. As we have argued, DSP screening by amplification and direct sequencing could increase by an order of magnitude the number of available polymorphic markers at any cloned locus. This technique encompasses and supercedes restriction enzyme-based screening since RFLPs and VNTRs may also be detected. As the only requirement for utilization of such markers is knowledge of a unique set of amplification primer sequences and of the polymorphism itself, publication of a polymorphism immediately makes it available to all readers. Hence, problems and delays associated with the physical transfer of plasmid DNAs between laboratories are avoided, and the costs of maintaining plasmid repositories will be ultimately reduced. In addition, rapid analysis of these polymorphic markers can be carried out on a large scale with the use of allele-specific oligonucleotide probes for direct hybridization to amplified DNA. Finally,

- 39 -

based on our experience with both strategies at the same locus, we found the expense and effort required to locate DSPs by either method to be comparable.

Treatment of Patients Having a Defective Rb Gene

In addition to screening, the invention includes polypeptide therapy for those individuals determined to contain a defective Rb allele, and who therefore are at risk of developing retinoblastoma.

To prevent the formation of retinoblastoma in these individuals, the Rb polypeptide is administered therapeutically in an amount sufficient to inhibit retinoblastoma tumor formation or growth (anti-retinoblastoma-forming amount). An anti-retinoblastoma-forming dosage of the Rb polypeptide is 1 to 500 µg/kilogram of body weight/day. The Rb protein can be administered by injection with a pharmacologically acceptable carrier, either alone or in combination with another agent. Acceptable pharmacological carriers are those which dissolve the Rb polypeptide or hold it in suspension, and which are not toxic to the extent of permanently harming the patient. Preferred are aqueous solutions of salts or non-ionic compounds such as sodium chloride or glucose, most preferably at an isotonic concentration. Other agents may be present provided that they do not interfere with the action of the Rb polypeptide. Those skilled in the art will know, or will be able to ascertain with no more than routine experimentation,

- 40 -

particular pharmacological carriers for this composition.

Rb polypeptide suitable for therapy can be prepared by any one of the following three conventional procedures. First, the Rb polypeptide can be produced by cloning the Rb cDNA from p4.7R into an appropriate mammalian expression vector, expressing the Rb gene product from this vector in an in vitro expression system, and isolating the Rb polypeptide from the medium or cells of the expression system. General expression vectors and systems are well known in the art.

Second, the Rb polypeptide can be produced using protein chemistry techniques, wherein the specific amino acid residues are joined together synthetically in the appropriate sequence.

Third, naturally occurring Rb protein can be isolated from total protein samples by affinity chromatography. Antibodies specific for the Rb protein are prepared by standard procedures (see below) and coupled to an inert matrix, which is then used to selectively bind the Rb proteins.

#### Immunodiagnosis of Retinoblastoma

This invention also includes methods for determining whether a particular tumor is the result of an Rb gene abnormality. Since osteosarcomas and certain undifferentiated tumors can result from detectable lesions in the Rb gene, immunodiagnosis can be used to aid in the diagnosis of such tumors.

- 41 -

In order to produce anti-Rb antibody, a rabbit is immunized with either naturally occurring Rb protein or Rb polypeptide produced as described above. The anti-Rb antibody generated is then labeled, e.g., radioactively, fluorescently, or with an enzyme such as alkaline phosphatase. The labeled antibody is used to determine whether human tumors are of defective Rb gene origin. This can be carried out using any conventional technique. For example, the tumor sample can be liquified and tested against the labeled antibody using a conventional ELISA (Enzyme-linked immunosorbent assay) format. Alternatively, human tissue samples (e.g., biopsy samples) can be tested for expression of the retinoblastoma protein by other immunological techniques, see e.g., I. Roitt, *Interaction of Antigen and Antibody, In Essential Immunology*, Fifth edition, Boston: Blackwell Scientific Publications, 1984, pp. 145-75.

Immune complexes will be detected in tumor samples which have antigens (e.g., retinoblastoma polypeptide) reactive with anti-Rb antibody. Tumors which lack these antigens presumptively have a defect (e.g., mutation or a deletion) in the retinoblastoma gene.

#### Deposits

Plasmids p2AR3.8 and p2AR0.9 were deposited on July 17, 1987 with the American Type Culture Collection, Rockville, Maryland, and assigned ATCC accession numbers 40,241 and 40,242, respectively.

- 42 -

The Applicants represent the ATCC is a depository affording permanence of the deposit and ready accessibility thereto by the public if a patent is granted. All restrictions on the availability to the public of the material so deposited will be irrevocably removed upon the granting of a patent. The material will be available during the pendency of the patent application to one determined by the Commissioner to be entitled thereto under 37 CFR 1.14 and 35 USC 122. The deposited material will be maintained with all the care necessary to keep it viable and uncontaminated for a period of at least five years after the most recent request for the furnishing of a sample of the deposited microorganism, and in any case, for a period of at least thirty (30) years after the date of deposit or for the enforceable life of the patent, whichever period is longer. Applicant acknowledges its duty to replace the deposit should the depository be unable to furnish a sample when requested due to the condition of the deposit.

Other embodiments are within the following claims.

08211

Table 4. Polymorphic sequences characterized and primer pairs used for PCR-amplification

Polymorphic Sequence	Allele Frequency	Amplification Primer Pair	Fragment Size	Location
RBI.2: TAAATAAGAICCTAACG TAAATAAGA CTTAACG	>95% < 5%	5'-AAGTGTAAATGGTTCTTAAG-3' 5'-TAGCAGGGTAATTTCCTC-3'	431 bp	124 bp from 5' end of exon 2
RBI.3: CAGAATTCGTTTCCTTT CAGAATTCAATTCCTTT	73% 27%	5'-TTCAAATATAGCCATCAGA-3' 5'-GCTTACACATGAATAGTGAGAG-3'	530 bp	43 bp from 3' end of exon 3
RBI.20: GATT(GIII) <sub>n</sub> CCTTT <sub>n=14-26</sub>	N/D	5'-AATTAAACAGGAGGTGGTGG-3' 5'-CTTGTAAATATGCCTCATTAAT-3'	550- 600 bp	54 bp from 3' end of exon 20
RBI.26: ATTTTTAAATCTGCAGT ATTTTTAAATCTGCAGT	85% 15%	5'-ATTCAAGTGAAGATACTTAAT-3' 5'-TAGTTCCCTTTGTAGTTCT-3'	683 bp	10 bp from 5' end of exon 26

Shown are the sequences and locations of the polymorphic sites and their immediate flanking regions. Also shown are the oligonucleotide primers used to amplify these sequences from human genomic DNA. Fragment size refers to the PCR-amplified product. Allele frequencies are based on analysis of the following numbers of individuals (of mixed North American descent): (RBI.2)-15; (RBI.3)-41; (RBI.20)-14; (RBI.26)-27. Accurate allele frequencies for RBI.20 have not yet been determined, as numerous alleles were found.

-44-

Table 3  
DNA sequence polymorphisms detected by direct sequencing

	Base Pairs Screened	Polymorphisms
Introns	2072	4
Exons	1640	0
Totals	3712	4

DNA sequence polymorphisms found by direct sequencing of 13 separate PCR-amplified regions from the human retinoblastoma locus. DNA samples from a minimum of 9 unrelated individuals were examined for all bases screened. Bases that could not be scored unambiguously were excluded from this tabulation.

- 45 -

Claims

1. Purified nucleic acid comprising a human retinoblastoma gene, or a fragment thereof comprising 15 or more bases, said nucleic acid being less than 100kb in size.
2. A vector comprising the nucleic acid of claim 1.
3. A cell transformed with DNA encoding retinoblastoma polypeptide or a fragment thereof.
4. The nucleic acid of claim 1, wherein said nucleic acid hybridizes specifically to said retinoblastoma gene under hybridizing conditions.
5. An isolated polypeptide encoded by the nucleic acid of claim 1.
6. An antibody produced to the polypeptide of claim 5.
7. An antibody produced to naturally occurring retinoblastoma polypeptide.
8. A method of detecting large deletions in the retinoblastoma gene of a human patient predisposing said patient to retinoblastoma, comprising the steps of:
  - hybridizing a nucleic acid sample from said patient with a probe specific for the retinoblastoma gene, and
  - determining the ability of said probe to hybridize to said nucleic acid,wherein lack of hybridization to said nucleic acid indicates the presence of a large deletion in said gene.
9. A method of detecting large deletions in the retinoblastoma gene of a human patient that may predisposing said patient to retinoblastoma, comprising the steps of:

- 46 -

generating nucleic acid fragments from a sample of said patient,

separating said fragments according to a determined physical property of said fragments,

hybridizing a probe specific for the retinoblastoma gene to said fragments,

detecting hybrids of said probe and said fragments, and comparing said hybrids to hybrids detected from the hybridization of said probe and separated nucleic acid fragments from a normal retinoblastoma gene,

wherein the absence of hybrids, or the smaller size of said hybrids from the sample of said patient is an indication of large deletions in the retinoblastoma gene of said patient.

10. The method of claim 8 or 9, wherein the probe specific for the retinoblastoma gene is the cloned DNA in p4.7R, or a fragment thereof.

11. The method of claim 9, wherein the physical property is molecular weight.

12. A method of detecting small deletions or point mutations in the retinoblastoma gene of a human patient predisposing said patient to retinoblastoma, comprising the steps of:

determining the nucleotide sequence of a retinoblastoma allele, or subregion thereof, from said patient, and

- 47 -

comparing said nucleotide sequence with the nucleotide sequence of a retinoblastoma allele or region thereof from a person not afflicted with retinoblastoma.

13. A method of detecting small deletions or point mutations in the retinoblastoma gene of a human patient predisposing said patient to retinoblastoma, comprising detecting mismatches between a nucleic acid sample from said patient and a detectable probe specific for the retinoblastoma gene from a person not afflicted with retinoblastoma, wherein mismatches are an indication of small deletions or mutations in the retinoblastoma gene of said patient.

14. A method of diagnosing predisposition of a human patient to retinoblastoma, comprising detecting the co-inheritance of retinoblastoma alleles of said patient with DNA polymorphisms in a pedigree analysis.

15. A method of detecting genetic polymorphisms, in the retinoblastoma gene of a human patient, that predispose said patient to retinoblastoma, comprising the steps of:

generating nucleic acid fragments from a sample of said patient,

separating said fragments according to a determined physical property of said fragments,

hybridizing a detectable nucleic acid probe capable of hybridizing to the wild type retinoblastoma gene to said fragments,

detecting hybrids of said probe and said fragments, and

- 48 -

comparing said hybrids to hybrids detected from the hybridization of said probe and separated nucleic acid fragments from a sample of a parent of said patient,

wherein the co-inheritance of specific genetic polymorphisms with the retinoblastoma gene is an indication of the predisposition of said patient to retinoblastoma.

16. The method of claim 15, wherein said physical property is molecular weight.

17. A method of treating a human patient having a defective retinoblastoma gene comprising administering to said patient an anti-retinoblastoma-forming amount of the retinoblastoma polypeptide.

18. A composition suitable for treating a human patient having a defective retinoblastoma gene, comprising retinoblastoma polypeptide and a pharmacologically acceptable carrier therefor.

19. A method of detecting the presence, in a tumor sample, of a protein the absence of which is associated with a neoplasm, said method comprising producing an antibody to said protein, contacting said antibody with said tumor sample, and detecting immune complexes as an indication of the presence in said tumor sample of said protein.

20. A method of detecting the presence of the retinoblastoma protein in a tumor sample from a human patient, comprising the steps of:

- 49 -

contacting said tumor sample with an antibody which specifically reacts with the retinoblastoma protein, and

determining whether immune complexes are formed with said antibody, the formation of said immune complexes being an indication that the tumor is not retinoblastoma and the absence of immune complexes indicating that the tumor is retinoblastoma.

21. The method of claim 20, wherein said antibody is a monoclonal antibody.

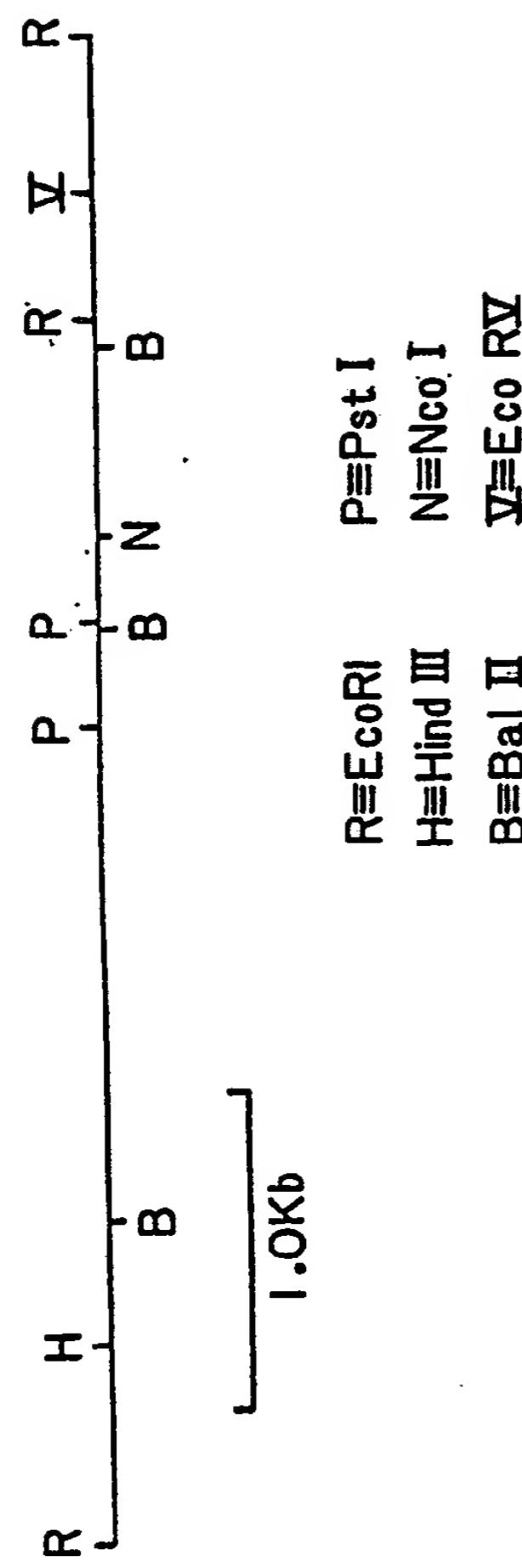


FIG. 1

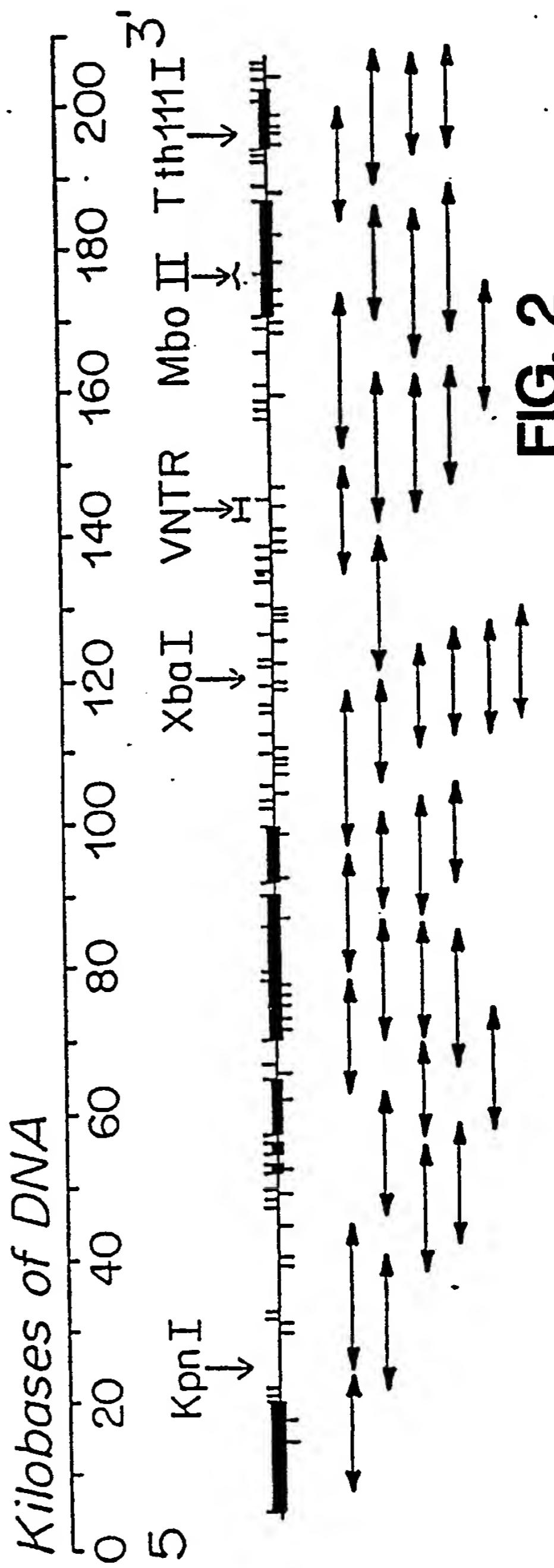
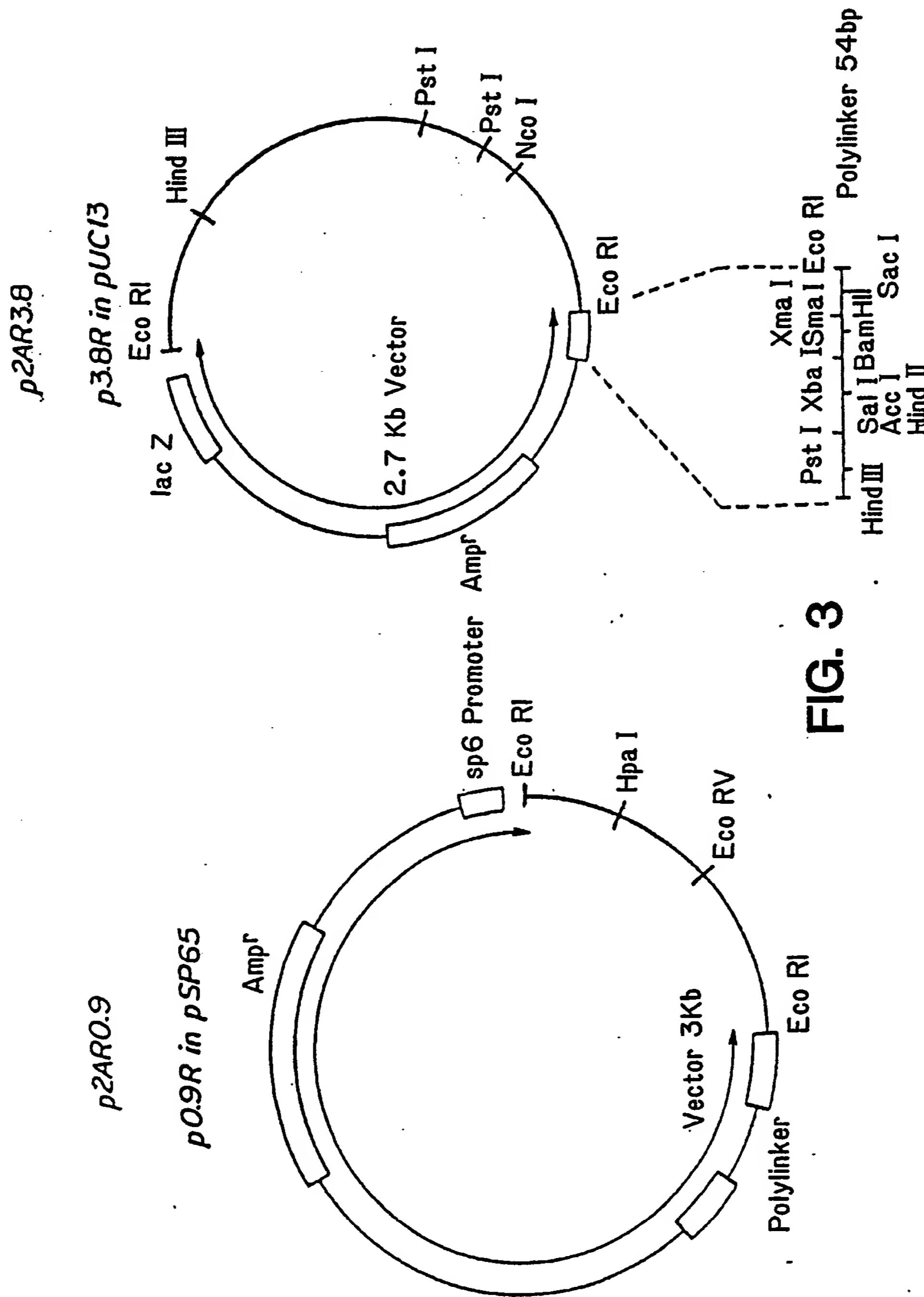


FIG. 2

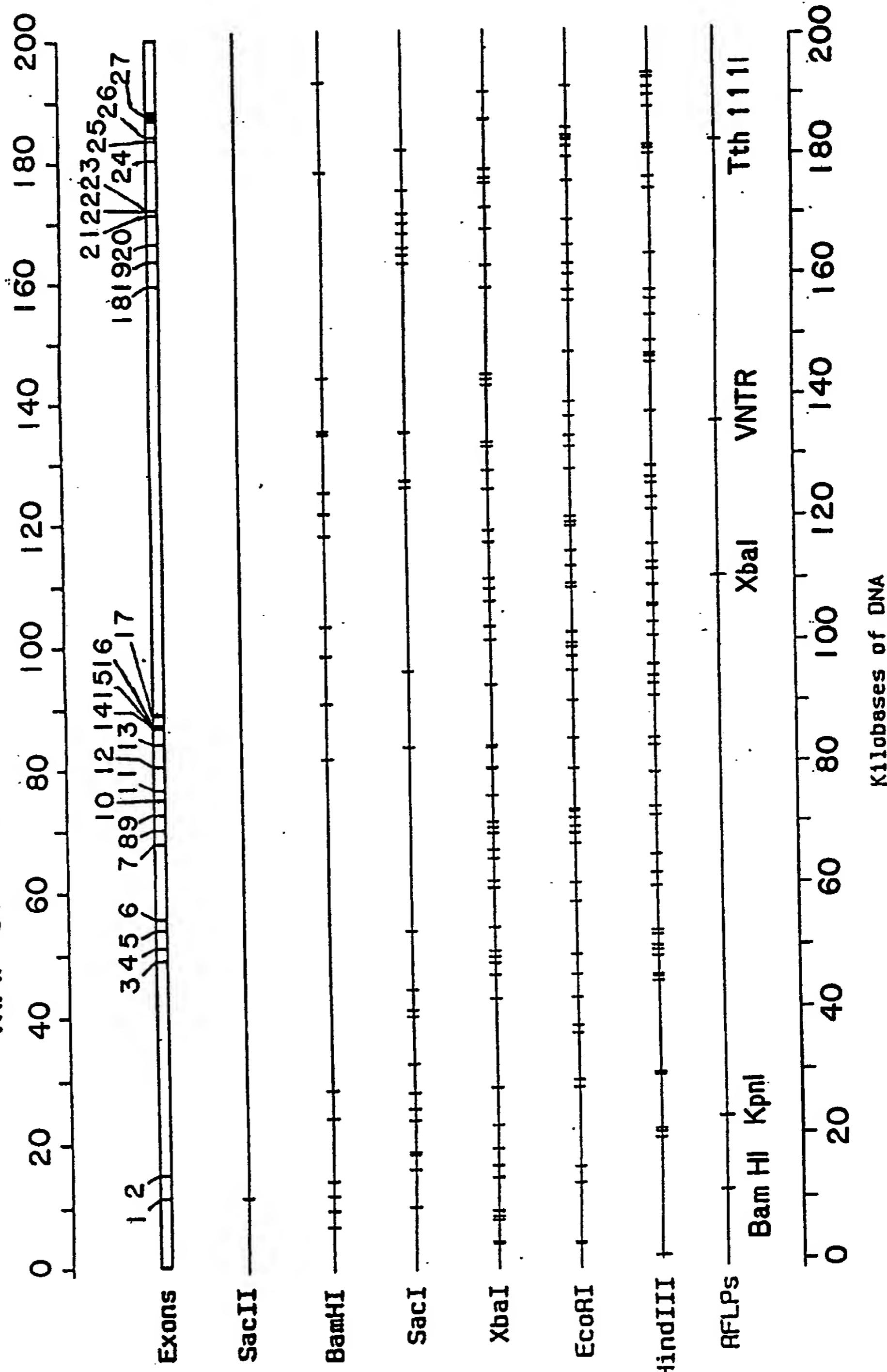
SUBSTITUTE SHEET

2/18

**FIG. 3****SUBSTITUTE SHEET**

3 / 18

**FIG. 4**  
**MAP OF THE RETINOBLASTOMA GENE**



**SUBSTITUTE SHEET**

4/18

## FIG. 5-1

GTCATGCCGCCAAAACCCCCGAAAAACGGCCGCCACCGCCGCCGCTGCCGCCGCCGGAACCCCCGGCACCGGCCGC...80  
 CCCCTCCTGAGGAGGGACCCAGAGCAGGACAGCGGCCGGAGACCTGCCTCTCGTCAGGCTTGAAGAAACAGAAG..160  
 AACCTGATTTACTGCATTATGTCAAGAAATAAGATACCAGATCATGTCAGAGAGAGAGCTTGGTTACTTGGGAGAAA..240  
 GTTTCATCTGGATGGAGTATTGGGAGGTTATATTCAAAGAAAAGGAACTGTGGGAATCTGTATCTTATTGCAGC..320  
 AGTTGACCTAGATGAG.....336

337	ATG TCG TTC ACT TTT ACT GAG CTA CAG AAA AAC ATA GAA ATC	378
	M S F T F T E L Q K N I E I	
379	AGT GTC CAT AAA TTC TTT AAC TTA CTA AAA GAA ATT GAT ACC	420
	S V H K F F N L L K E I D T	
421	AGT ACC AAA GTT GAT AAT GCT ATG TCA AGA CTG TTG AAG AAG	462
	S T K V D N A M S R L L K K	
463	TAT GAT GTA TTG TTT GCA CTC TTC AGC AAA TTG GAA AGG ACA	504
	Y D V L F A L F S K L E R T	
505	TGT GAA CTT ATA TAT TTG ACA CAA CCC AGC AGT TCG ATA TCT	546
	C E L I Y L T Q P S S S I S	
547	ACT GAA ATA AAT TCT GCA TTG GTG CTA AAA GTT TCT TGG ATC	588
	T E I N S A L V L K V S W I	
589	ACA TTT TTA TTA GCT AAA GGG GAA GTA TTA CAA ATG GAA GAT	630
	T F L L A K G E V L Q M E D	
631	GAT CTG GTG ATT TCA TTT CAG TTA ATG CTA TGT GTC CTT GAC	672
	D L V I S F Q L M L C V L D	
673	TAT TTT ATT AAA CTC TCA CCT CCC ATG TTG CTC AAA GAA CCA	714
	Y F I K L S P P M L L K E P	
715	TAT AAA ACA GCT GTT ATA CCC ATT AAT GGT TCA CCT CGA ACA	756
	Y K T A V I P I N G S P R T	
757	CCC AGG CGA GGT CAG AAC AGG AGT GCA CGG ATA GCA AAA CAA	798
	P R R G Q N R S A R I A K Q	
799	CTA GAA AAT GAT ACA AGA ATT ATT GAA GTT CTC TGT AAA GAA	840
	L E N D T R I I E V L C K E	
841	CAT GAA TGT AAT ATA GAT GAG GTG AAA AAT GTT TAT TTC AAA	882
	H E C N I D E V K N V Y F K	
883	AAT TTT ATA CCT TTT ATG AAT TCT CTT GGA CTT GTA ACA TCT	924
	N F I P F M N S L G L V T S	
925	AAT GGA CTT CCA GAG GTT GAA AAT CTT TCT AAA CGA TAC GAA	966
	N G L P E V E N L S K R Y E	
967	GAA ATT TAT CTT AAA AAT AAA GAT CTA GAT GCA AGA TTA TTT	1008
	E I Y L K N K D L O A R L F	
1009	TTG GAT CAT GAT AAA ACT CTT CAG ACT GAT TCT ATA GAC AGT	1050
	L D H D K T L Q T D S I D S	
1051	TTT GAA ACA CAG AGA ACA CCA CGA AAA AGT AAC CTT GAT GAA	1092
	F E T Q R T P R K S N L D E	
1093	GAG GTG AAT GTA ATT CCT CCA CAC ACT CCA GTT AGG ACT GTT	1134
	E V N V I P P H T P V R T V	
1135	ATG AAC ACT ATC CAA CAA TTA ATG ATG ATT TTA AAT TCA GCA	1176
	M N T I Q Q L M M I L N S A	
1177	AGT GAT CAA CCT TCA GAA AAT CTG ATT TCC TAT TTT AAC AAC	1218
	S D Q P S E N L I S Y F N N	

SUBSTITUTE S...

5/18

**FIG. 5-2**

1219	TGC ACA GTG AAT CCA AAA GAA AGT ATA CTG AAA AGA GTG AAG C T V N P K E S I L K R V K	1260
1261	GAT ATA GGA TAC ATC TTT AAA GAG AAA TTT GCT AAA GCT GTG D I G Y I F K E K F A K A V	1302
1303	GGA CAG GGT TGT GTC GAA ATT GGA TCA CAG CGA TAC AAA CTT G Q G C V E I G S Q R Y K L	1344
1345	GGA GTT CGC TTG TAT TAC CGA GTA ATG GAA TCC ATG CTT AAA G V R L Y Y R V M E S M L K	1386
1387	TCA GAA GAA GAA CGA TTA TCC ATT CAA AAT TTT AGC AAA CTT S E E E R L S I Q N F S K L	1428
1429	CTG AAT GAC AAC ATT TTT CAT ATG TCT TTA TTG GCG TGC GCT L N D N I F H M S L L A C A	1470
1471	CTT GAG GTT GTA ATG GCC ACA TAT AGC AGA AGT ACA TCT CAG L E V V M A T Y S R S T S Q	1512
1513	AAT CTT GAT TCT GGA ACA GAT TTG TCT TTC CCA TGG ATT CTG N L D S G T O L S F P W I ; L	1554
1555	AAT GTG CTT AAT TTA AAA GCC TTT GAT TTT TAC AAA GTG ATC N V L N L K A F D F Y K V I	1596
1597	GAA AGT TTT ATC AAA GCA GAA GGC AAC TTG ACA AGA GAA ATG E S F I K A E G N L T R E M	1638
1639	ATA AAA CAT TTA GAA CGA TGT GAA CAT CGA ATC ATG GAA TCC I K H L E R L E H R I M E S	1680
1681	CTT GCA TGG CTC TCA GAT TCA CCT TTA TTT GAT CTT ATT AAA L A W L S D S P L F D L I K	1722
1723	CAA TCA AAG GAC CGA GAA GGA CCA ACT GAT CAC CTT GAA TCT Q S K D R E G P T O H L E S	1764
1765	GCT TGT CCT CTT AAT CTT CCT CTC CAG AAT AAT CAC ACT GCA A C P L N L P L Q N N H T A	1806
1807	GCA GAT ATG TAT CTT TCT CCT GTA AGA TCT CCA AAG AAA AAA A D M Y L S P V R S P K K K	1848
1849	GGT TCA ACT ACG CGT GTA AAT TCT ACT GCA AAT GCA GAG ACA G S T T R V N S T A N A E T	1890
1891	CAA GCA ACC TCA GCC TTC CAG ACC CAG AAG CCA TTG AAA TCT Q A T S A F Q T Q K P L K S	1932
1933	ACC TCT CTT TCA CTG TTT TAT AAA AAA GTG TAT CGG CTA GCC T S L S L F Y K K V Y R L A	1974
1975	TAT CTC CGG CTA AAT ACA CTT TGT GAA CGC CTT CTG TCT GAG Y L R L N T L C E R L L S E	2016
2017	CAC CCA GAA TTA GAA CAT ATC ATC TGG ACC CTT TTC CAG CAC H P E L E H I I W T L F Q H	2058
2059	ACC CTG CAG AAT GAG TAT GAA CTC ATG AGA GAC AGG CAT TTG T L Q N E Y E L M R D R H L	2100
2101	GAC CAA ATT ATG ATG TGT TCC ATG TAT GGC ATA TGC AAA GTG D Q I M M C S M Y G I C K V	2142
2143	AAG AAT ATA GAC CTT AAA TTC AAA ATC ATT GTA ACA GCA TAC K N I D L K F K I I V T A Y	2184

6/18

**FIG. 5-3**

2185	AAG GAT CTT CCT CAT GCT GTT CAG GAG ACA TTC AAA CGT GTT	2226
	K D L P H A V Q E T F K R V	
2227	TTG ATC AAA GAA GAG GAG TAT GAT TCT ATT ATA GTA TTC TAT	2268
	L I K E E Y D S I I V F Y	
2269	AAC TCG GTC TTC ATG CAG AGA CTG AAA ACA AAT ATT TTG CAG	2310
	N S V F M Q R L K T N I L Q	
2311	TAT GCT TCC ACC AGG CCC CCT ACC TTG TCA CCA ATA CCT CAC	2352
	Y A S T R P P T L S P I P H	
2353	ATT CCT CGA AGC CCT TAC AAG TTT CCT AGT TCA CCC TTA CGG	2394
	I P R S P Y K F P S S P L R	
2395	ATT CCT GGA GGG AAC ATC TAT ATT TCA CCC CTG AAG AGT CCA	2436
	I P G G N I Y I S P L K S P	
2437	TAT AAA ATT TCA GAA GGT CTG CCA ACA CCA ACA AAA ATG ACT	2478
	Y K I S E G L P T P T K M T	
2479	CCA AGA TCA AGA ATC TTA GTA TCA ATT GGT GAA TCA TTC GGG	2520
	P R S R I L V S I G E S F G	
2521	ACT TCT GAG AAG TTC CAG AAA ATA AAT CAG ATG GTA TGT AAC	2562
	T S E K F Q K I N Q M V C N	
2563	AGC GAC CGT GTG CTC AAA AGA AGT GCT GAA GGA AGC AAC CCT	2604
	S O R V L K R S A E G S N P	
2605	CCT AAA CCA CTG AAA AAA CTA CGC TTT CAT ATT GAA GGA TCA	2646
	P K P L K K L R F D I E G S	
2647	GAT GAA GCA GAT GGA AGT AAA CAT CTC CCA GGA GAG TCC AAA	2688
	D E A D G S K H L P G E S K	
2689	TTT CAG CAG AAA CTG GCA GAA ATG ACT TCT ACT CGA ACA CGA	2730
	F Q Q K L A E M T S T R T R	
2731	ATG CAA AAG CAG AAA ATG AAT GAT AGC ATG GAT ACC TCA AAC	2772
	M Q K Q K M N D S M D T S N	
2773	AAG GAA GAG AAA	2784
	K E E K	
TGAGGATCTCAGGACCTTGGTGGACACTGTGTACACCTCTGGATTCTGCTCTCACAGATGTGACTGTATAACCTTCC		
	CAGGTTCTGTTATGGCCACATTTAATATCTTCAGCTTTTGTGGATATAAAATGTGCAGATGCAATTGTTGGGTGA	2944
	TTCTAACGCCACTGAAATGTTAGTCATTGTTATTTATACAAGATTGAAAATCTGTGTAAATCCTGCCATTAAAAAGT	3024
	TGTAGCAGATTGTTCTCTCAAAGTAAAATTGCTGTGCTTATGGATAGTAAGAATGGCCCTAGAGTGGGAGTCCTG	3104
	ATAACCCAGGCCTGCTGACTACTTGCCTCTTTGTAGCATATAGGTGATGTTGCTCTTGTATTAATTATAT	3184
	GTATATTTTTAATTAAACATGAACACCCCTAGAAAATGTGTCCATCTATCTTCAAATGCAATTGATTGACTGCC	3264
	ATTCACCAAAATTATCCTGAACTCTTCTGCAAAAATGGATATTATTAGAAAATTAGAAAAAAATTACTAATTTCACATT	3344
	AGATTTATTTACTATTGGAATCTGATATACTGTGTGCTTGTATAAAATTGCTTTAATTAAATAAAAGCTGGA	3424
	AGCAAAGTATAACCATATGATACTATCATACTACTGAAACAGATTCTACCTCAGAATGAAAAGAACTTACTGATTAT	3504
	TTTCTCATCCAACCTATGTTTAAATGAGGATTATTGATAGTACTCTGGTTTATACCATTCACTGAATT	3584
	ATAAAGTACCCATCTAGTACTGAAAAAGTAAAGTGTCTGCCAGATCTTAGGTATAGAGGCCCTAACACAGTATATCC	3664
	CAAGTGCACTTCTAATGTTCTGGGTCTGAAGAATTAAGATACAAATTATTTACTCCATAAACAGACTGTTAATTA	3744
	TAGGAGCCTTAATTTTTCTAGAGATTGTCTAATTGCACTCTCAAAATTATTCTGCCCTCTTAATTGGGAAGGT	3824
	TTGTGTTTCTGGAATGGTACATGTCTTCATGTATCTTGAACCTGGCAATTGTCTATTATCTTATTAA	3904
	GTCAGTATGGCTAACACTGGCATGTTCAAAGCCACATTATTCAGTCCAAAATTACAAGTAATCAAGGGTCAATTGG	3984
	GTTAGGCATTAATGTTCTATCTGATTTGTGCAAAGCTTCAAATTAAAACAGCTGCAATTAGAAAAGAGGCCTCTC	4064
	CCCTCCCCACACCTAAAGGTGATTTAAACTATCTGTGTGATTAACCTATTAGAGATGCTGTAACCTAAAATAGGG	4144
	ATATTAAAGGTAGCTCAGCTAGCTTTAGGAAAATCACTTGTCTAACTCAGAATTATTTAAAAAGAAATCTGGTCT	4224
	TGTTAGAAAACAAAATTATTTATTTGTGCTCATTTAAGTTCAAACCTACTATTGACAGTTATTGATAACAATGACA	4304
	CTAGAAAACCTGACTCCATTCTCATCATTGTTCTGCATGAATATCATACAAATCAGTTAGTTAGGTCAAGGGCTTAC	4384
	TATTTCTGGGTCTTTGCTACTAAGTTCACATTAGAATTAGTGCCTAGAATTAGTCAAGGACTTCAGAGATCGTGTATTGAGA	4464
	TTTCTTAAATAATGCTTCAGATATTATTGCTTATTGCTTTGTATTGGTTAAAACGTACATTAAAATTGCTATGT	4544
	TACTATTTCTACAATTAAAGTTGTCTATTAAAATAATTAGTTGTAG.....	4597

7/18

## FIG. 6-1

start of cDNA sequence

<u>SacII</u>	aaggagggag agtggngtcc ngnngagggt gcactagcca gatattctgc ggggccc gag	060
<u>EXON 1</u>	agtcttcct atcagacccc gggatagggta gaggCCCAC AGTCACCCAC CAGACTCTT	120
	GTATAGCCCC GTTAAGTGCA CCCCGGCCCTG GAGGGGTGG TTCTGGTAG AAGCACGTCC	180
	GGGCCGCGCC GGATGCCTCC TGGAAAGGCC CTGGACCCAC GCCAGGTTTC CCAGTTAAT	240
	TCCTCATGAC TTAGNGTCCC AGCCNGCGA CCGACCAGCG CCCCAGTTCC CCACAGACGC	300
	<u>SacII</u> CGGCGNNNC GGGAGCCTGC GGACGTGAGC GCGGGCGGAA GTGACGTTTT CGCGGGTTGG	360
	<u>EXON 1</u> ACGCGGCGCT CAGTTGCCGG GCGGGGGAGG GCGCGTCCGG TTTTCTCAG GGGACGTTGA	420
	MetProPr (3)	
<u>SacII</u>	AATTATTTT GTAACGGGAG TCGGGAGAGG ACGGGGCGTG CCCCGCGTGC GCGCGCGTGC	480
	TCCTCCCCGG CGCTCCTCCA CAGCTCGCTG GCTCCCGCC CGGAAAGGCG TCATGCCGCC	540
	TCTCGTCAGG TGAGCGAGCA GAGCGCGTCN CTCACGCGGG AAGGGCGCCC CGGGTGTGCG	720
	(46) oLeuValAr TAGGGCGGGC GCAAGGCGGC TCGGGGGGA CCCGTCTCG CCAGGGCCG GGTCCcgGNG	780
	GGAGGAGGGG CCCTCCCTGC CCCCGCCAC GGcggaGCGT CTGCAGAATG GTGACAGGAT	840
	TCTGGTTCT TGGCGAGGG GTCTCGGCTT CAACTTGACA GGTGTGGGC GGGTgggct	900
	agnntctga gcaagtgac aggtgcagtt ccctttgtg agnetcggan ncagaggntc	960
	gttgcgagcg tncatcagac aaaaaaatga aaaataaaa tacaaaaa	1008
	--2.9 kb--	
<u>EXON 2</u>	ccccaaacage ttttagctatt acatttactt tccttcacag aagtgtttt ctgtttgaa	060
	gatatttgc ttaccatgca agcaaattt tttcaactgtg tggtatcctt attttggaaat	120
	gaccatgaaa aagataatca tatgnnnnaaa tttgaagtgt aatgttttc taagataaaa	180
	taagatctt AAGTATTIAA TAATGTTCTT TTTCACAGTA GTGTTATGTG CAAACTATTG	240
	AAACAAAGTAT GTACTGAATC AATTGATTATAAAGATATG CCAATTATAT GATTATTTTC	300
	ATTGGTAGG CTTGAGTTTG AAGAACAGA AGAACCTGAT TTTACTGCAT TATGTCAGAA	360
	g LeuGluPheG luGluThrGl uGluProAsp PheActAlaL euCysGlnLy	(63)
	ATTAAAGATA CCAGATCATG TCAGAGAGAG AGCTGGTTA ACTTGGGAGA AAGTTTCATC	420
	sLeuLysIle ProAspHisV alArgGluAr gAlaTrpLeu ThrTrpGluL ysValSerSe	(83)
	TGTGGATGGA GTATTGGTAA GGATTTCTT AAAACGTTTT GAAATTTTT TTTCTCATT	480
	rValAspGly ValLeu TAAAACCAAC TTCAAATCAC TATACAAAAA TTGAAAGATA GAAAATATA AAGACAATAA	540
	AAGctaataa taattccatt acccagagga aattcaccc tgctaacatt aaaaatgttt	600
	gaggccgggc acgtggttca tgctgttaat cctaccatt tgggaggctg aggccagggtgg	660
<u>SacI</u>	atgcctgag ctcaggagtt cgagaccgc ctggcaaca tggt	704
	--33 kb--	
<u>EXON 3</u>	ctatttgaga tgactgaccc ctaaagttcc acaataacta tttaattttt tatctttcta	060
	atactttttt gccttataat staaaattt aatgtttgtt attagtgtga aatgaaatcc	120
	tttcaaatat atgccatcg aaggatgtgt tacaaatata cagtATTACA AACATTATT	180
	TTGTATGCTG AATAAGAAAA AATCAGTTAT AATACAGTTT TAACATAGTA TCCAGTGTGT	240
	GAATTATTAA ATGAAATATT TGATCTTAT TTTTGTCC AGGGAGGTAA TATTCAAAAG	300
	GlyGlyTy rIleGlnLys	(94)
	AAAAAGGAAC TGTGGGAAT CTGTATCTT ATTGCAGCAG TTGACCTAGA TGAGATGTCC	360
	LysLysGluL euTrpGlyI1 eCysIlePhe IleAlaAlaV alAspLeuAs pGluMetSer	(114)
	TTCACCTTAA CTGAGCTACA GAAAACATA GAAATCAGGT AAAGTTCTT GTATAAAATAT	420
	PheThrPheT hrGluLeuGl nLysAsnIle GluIleSe	(127)
<u>EcoRI</u>	AAGCCTCTGC CATAAAAGGA AACGAATTCT GGATTTCTT CTCAATAGAC TTTTGTGAAT	480
	TAGTGAGAAA TGCTAAAATA AAGTAAAACA AAAAGAACTT GGACCAAATA GTGAACTGCC	540
	ATTCTCTCAT GGAGCCGTTA TGAAAGTGTG TTTATGCTGT ATTCTTTAA GAGGTAGCAG	600
	TTTGTGTCTT GGAAAAATTTC CATTGTGTG TCTCACTATT CATGTGTAAAG C	651
	--1.6 kb--	

8 / 18

**FIG. 6-2**

gcataggat atagataata gaggtgttaag ttgaaggcta attatttttca caaaaagctaa 060  
 ttccttccaa aggatatagt agtgatttga tgttagagctg ataatcTTT GAATTGAAAT 120  
 ATCTATGATT TGAAAACGAA ATAACACAAA TTTTAAGGT TACTGATTIA CTTTTTCTA 180  
 EXON 4 TTCTTCCTT TGTAGTGTCC ATAAATTCTT TAACCTACTA AAAGAAATTG ATACCAGTAC 240  
 rValH isLysPhePh eAsnLeuLeu LysGluIleA spThrSerTh (142)  
 CAAAGTTGAT AATGCTATGT CAAGACTGTT GAAGAAGTAT GATGTATTGT TTGCACTCTT 300  
 rLysValAsp AsnAlaMetS erArgLeuLeu uLysLysTyr AspValLeuP heAlaLeuPh (162)  
 CAGCAAATTG GAAAGGTAAA GTAAACATT TATTAGGGTT ACACCTGAT TTTTATGTC 360  
 eSerLysLeu GluAr (167)  
 ATTGTTCACA ATTAGATTCT GGGATTATT TAACACATT AGTAAAGTTA GTAAGTATTA 420  
 ATTCTTAgac ttgtccctt taatgttagc tcattaattc ttatgtttt tatttateca 480  
 gtaatatgca ttctgaatgc ttctggaaa attaaccgtt attatccttt catgtctcca 540  
 tttgtttca aaacttagct tatcgagtat 570  
 --2.1 kb--  
 gagatattta aagagnaact ttactaacct taggtggatc agctgggtgt tttctatctt 060  
 atttataacct tttttttgaa GACTAATTGA GAGGATTAAC TGTAATTATA TATTAAAGTG 120  
 ATGTGAGATG TCATAAAATTG GGAAAATCTA CTTGAACCTT GTTTATAAT GCTATATATT 180  
 TTTTGTTTT AAAATATATA CTTCTAAAAA GAAGATGAAT AAAGCATGAG AAAACTACTA 240  
 EXON 5 TGACTTCTAA ATTACGAAAA AATGTTAAAA AGTCATAATG TTTTCTTT CAGGACATGT 300  
 gThrCys (169)  
 GAACTTATAT ATTGACACA ACCCAGGAGT TCGTAAGTAG TTCACAGAAT GTTATTITTC 360  
 GluLeuIleT yrLeuThrG1 nProSerSer Se (180)  
 ACTTAAAAAA AAAGATTTT ATGGAATAAT CTCAAACATC TTGATAGTTA GGGTTAGTTT 420  
 GATCGATTAT AGCAGGCTAC Ttcataaattt aagcccatag atttaagtcc tgtgttagatt 480  
 atttatcttc tcacaaagaa aatagtataa aatacatgcc ttgtactaca aagaagaact 540  
 aataaggtgg aattgattca ggacagcata tcaccaactc tgagaaaaat gcaacaaatg 600  
 caaattcatt gactaa 616  
 --1.4 kb--  
 aaatggactg cattctatata tgcatttaac taaggtcatt ttttttttaa tGCACAAAAA 060  
 GAAACACCCA AAAGATATAT CTGGAAAATC TTCTTCAGT GATACATT TTCTGTTTT 120  
 EXON 6 TTTCTGCTT CTATTGTTT AATAGGATAAT CTACTGAAAT AAATTCTGCA TTGGTGCTAA 180  
 rIleS erThrGluII eAsnSerAla LeuValLeuL (192)  
 AAGTTCTTG GATCACATTT TTATTAGCTA AAGGTAAAGTT CATTATATT ATTAAATGCT 240  
 ysValSerTr pIleThrPhe LeuLeuAlaL ysG (203)  
 AATATTCAA ATGTAATAAT TAAATTGGCA TTCCTTGGA CTAAATTCCC CAATTTTAT 300  
 TGAGTAATGT ACTCCTccct cattctctgc ttggcttatt aactgttagc aagttcctat 360  
 aattctggta cttagaaacaa ccttggaaat gctttattta atntttgttt ctaatattcc 420  
 atcttccctc cttt 434  
 --11.5 kb--  
 ttatagtgat ttttagacat aaagaattaa ttataacaga aatagcttaa atgtaaaatt 060  
 ctcagagttag agcttaacac ttgatttata attccataac ttacatatt tCTATTTAC 120  
 ATATTTATA CCTTTAAAAA CAGATTTTTT TTTTTTTAC AAAAAAAAGA AAGAAAATCT 180  
 TTACCATGCT GATAGTGATT GTGAATGAA TAAATTATG GATATACTCT ACCCTGGAT 240  
 TTTCTCTCAT ACAAAAGATCT GAATCTCTAA CTTCTTTAA AAATGTACAT TTTTTTTCA. 300  
 EXON 7 GGGGAAGTAT TACAAATGGA AGATGATCTG GTGATTICAT TTCAAGTTAAT GCTATGTGTC 360  
 1yGluVall euGlnMetG1 uAspAspLeu ValIleSerP heGlnLeuMe tLeuCysVal (222)  
 CTGACTATT TTATTAAACT CTCACCTCCC ATGTTGCTCA AAGAACCTAA TAGTAAGTAT 420  
 LeuAspTyrP heIleLysLe uSerProPro MetLeuLeuL ysGluProTy rL (240)  
 TTAATTATG CCCCTTTAC TTTCTCATTC AGCAGTTGCT TATTGAATGT CTAGTGGGT 480  
 CCAAACATGG TTCTAAGGCT GACAGGATGA TAAAAAAATAA ATCAAGacatg gactttggccc 540  
 ataagtagtg taagttatag aaggaaagat aagacatgga aacaatgat tagagtata 600  
 ggttagaaagt ggtttcggtt caaaatacaa caaatggagg tttgggagac aagaag 656  
 --1.8 kb--

9 / 18

**FIG. 6-3**

10/18

## FIG. 6-4

aatactgaac aacttggta tcaataccnc cagggagaag catctgactt tcactttaa 060  
 aaaaagactt aatgatttgt atacctttt gtcataaaca taatggaaag agaccacaa 120  
 ttaaaaagng tagtgAAAGG TATTTATT AAGCAGCAGC TGGTCATCT ATTTCTATC 180  
 CTATCTATTA TTGAGTTATC ATTTATATG ATTTATGAG ACAACAGAAG CATTATACTG 240  
 CTTTTTGAT GCATAAAGCA CAAATTGTAA ATTTCACTGA TGTGAATGAC TTCACTTATT 300  
 EXON 11 GTTATTAGT TTGAAACAC AGAGAACACC ACGAAAAAGT AACCTTGATG AAGAGGTGAA 360  
 r PheGluThrG llnArgThrPr oArgLysSer AsnLeuAspG luGluValAs (367)  
 TGTAATTCCCT CCACACACTC CAGTTAGGTA TGAATTTC TACTTTAAT TATATTATAA 420  
 nValIlePro ProHisThrP roValAr (326)  
 TTTGTTATT CATGGCTTA TAGTGTTC GATTGTTCA CGTTCTTA TGTATTCTATA 480  
 CATACTGTA AGAAATATAT ATTGAAGGCC AGGTGTGGTG GATCACACCT GTAATCCCAG 540  
 CACTTTGGGA GGCCAAGGCG GGCAGATCAC CTGAGGTTAG GAGTTTGaga ccggcttgcc 600  
 caacatggtg aaaccccgtc tctactagaa atacaaaaat tagctggggg tggtggtgtg 660  
 tgctgtaat ccagctgtc 680

--3.2 kb--

caataccatt ttgttgcag ttatatagtt ctccaaaaa taatgccACT ATTTTATTGA 060  
 TATGTAGTT TATTAGTAAA TAAGTATATC TGTTCTATAA CTATAAACTT ATTGATTGTG 120  
 AATACATATT TTCTAAAGA TTIAAGTAAA ATGTAATTTC TTATAAAACCA CAGTCTTATT 180  
 TGAGGGAAATG TAGAGACAAG TGGGAGGCAG TGTATTGAA GATACATTAA ACTTGGGAGA 240  
 TTGAAAACAT TTCAATTCTT CTTTTTTCT CCCTTCATTG CTAAACACAT TTCTCTATT 300  
 EXON 12 TTATCCCCTC TAGGACTGTT ATGAACACTA TCCAACAATT AATGATGATT TTAAATTCA (392)  
 gThrVal MetAsnThrI leGlnGlnLe uMetMetIle LeuAsnSerA  
 CAAGTGATCA ACCTTGAGAA AATCTGATT CCTATTAA CGTAAGCCAT ATATGAAACA 420  
 laSerAspG1 nProSerGlu AsnLeuIleS erTyrPheAs n (405)  
 TTATTATTG TAATATCTTG GCAAAGAAC TTGAAATTAA AAGTTAAAGT ACTGAGTTCT 480  
 TTTAAAATA CTAATCTCCT ATCTAACATG TAGTTATCCA TAATCTTTTC TTGCTTTTT 540  
 AATCTTACAA ATTATATATT ATTAGTAGTA TTGTTTTATT TATACAGTGT TATTAAAAC 600  
 ATTTTATGT TTACCTATT GCCTTgtca ccattcttcc ttcaactta tgcctcactt 660  
 ctgagataat tttttttct tcagatatac ctttgataaa ttac 704

--3.1 kb--

aaaatttaga taatagggtt ttttagttgt actgttagtat ttttgctcg attaacatcc 060  
 AGTGAATGA TATTGTCAG TAGTTGTGGT TACCTAGTTA TTATGGAAGT 120  
 GTTCCACAT TTTATGAAC AATTAAAAA GTCATATATT ATGGAGCAGA AAATATTAAT 180  
 EXON 13 TCTGATTACA CAGTATCCTC GACATTGATT TCTGTTTTA CCTCCTAAAG AACTGCACAG 240  
 AsnCysThrV (409)  
 TGAATCCAAA AGAAAGTATA CTGAAAAGAG TGAAGGATAT AGGATACATC TTTAAAGAGA 300  
 alAsnProLy sGluSerIle LeuLysArgV allysAspII eGlyTyrIle PheLysGluL (429)  
 EcoRI AATTGCTAA AGCTGTGGGA CAGGGTTGTG TCGAAATTGG ATCACAGGTA ACTTGAATTG 360  
 ysPheAlaLy sAlaValGly GlnGlyCysV alGluIleG1 ySerGln (444)  
 ATTGTAATTG GTGGTACTAT AGAGTAATAA TATTAAGC AGCATCTTC CAGTTGTAT 420  
 AAATACTCTA ACAGTATTG TCTAGTAGTA TAAAATCTG TCAGATACTA TATCCCTGCT 480  
 GCCTGTGTAT GCTGCTATT ATGGAAACTT TATGGAAAAC TACCTCCCAC cccattataa 540  
 aaactatgta ataaaggaac acatagccat tgtagsaatt ttng 584

--1.8 kb--

11/18

**FIG. 6-5**

--1.0 kb--

gaatgttaat caccactaa tacttaagt gtgagttta gacaagctng ctttctgttt 060  
 gtcttggcgg ccataattgt aagaagggtg AGAAGTATGT TTTAAGAAAA GGCTTTTAA 120  
 AAAATTTCAG TAATTCAG CTGGGTATAG TGGTACATGC CTATAATCCC AGCCTCTTGG 180  
 GAGGCCAAAG CAGGAGGATC TCTTGAGCC AGGAGTGTGA AGGCCAGCCT GGGCAAAACA 240  
 GTGAGACTCC ATCTCAAAAAA AAAAAAAA TTTCTATAATT GTGATTTCT AAAATAGCAG 300  
 EXON 14 GCTCTTATT TTCTTTGT TTGTTGTAG CGATACAAAC TTGGAGTTCG CTTGTATTAC 360  
 ArgTyrLysL euGlyValAr gLeuTyrTyr (454)  
 CGAGTAATGG AATCCATGCT TAAATCAGTA AGTIAAAAAC AATAATAAAA AAATTCANC 420  
 ArgValMetG luSerMetLe uLysSer (463)  
 CGGGCGCGGT GGCTCACGCC TGCAATCCCA GCACCTTGGG AGGCCGAGGT GGGCAGATCA 480  
 GGAGGTCAAG GCATCAAGAT CATCCTGGCC AAAATGGTGA AACCTGTCT CTACTAAAAG 540  
 TACAAAAATT AGCTGGGCGT GGTGGTGTAG ACCTGTAGTC CCAGCTACTT GGCAGGCTGA 600  
 GGCAGGAGAA TCCCTTGAAC CACGGAGGTG GAGGTGGAG TGAGCCAAGA TTGTGCCATT 660  
 TCACCCCAGC CTGGCAACAG AGCAAGACAC CATCTAAAAA AAAAAAAA AAAAAAAA 720  
 ATTCAATGCT GACACAAATA AGGTTGAAT TAAACAAACTT CTTTTTTTTT TTTTAAATTA 780  
 EXON 15 TCTGTTTCAG GAAGAAGAAC GATTATCCAT TCAAAATTG AGGTAAATTG TTIACTTTA 840  
 GluGluGluA rgLeuSerII eGlnAsnPhe Se (474)  
 GTAAAAAAATT TTTTCTTT TATAGAAGTA AGTATTTAT AATCTTTTTT TTTTCTTTT 900  
 EXON 16 AGCAAACCTTC TGAATGACAA CATTTCAT ATGCTTTAT TGGCGTGCAG TCTTGAGGTT 960  
 rLysLeuL euAsnAspAs nIlePheHis MetSerLeuL euAlaCysAl aLeuGluVal (493)  
 GTAATGGCCA CATATAGGAG TAAGTTAAAT TTTCTAAAT AAACACTTTT GTTCAATTAA 1020  
 ValMetAlaT hrTyrSerA (500)  
 AAGTTAAAAT GTGGTGTGTT TCTTGGTGC GGGGAGAGGG ATAGTGTGAG GTTAAGGAGA 1080  
 AGGAATGCTT ATTTAGATC ACTATATACT. GAAGAATGTA ATTGGTCATT ATAAGCCATT 1140  
 TAAGAGGCCTT ATTTGAGTTA TTTGAggcac tcttggttat aatatttcac taggcttctc 1200  
 ttctgagtat actggtatac tgaatccaaa aaaggtactt ttccgaaatc cctccgaaga 1260  
 cctttgagat tgttagagtgc 1280  
 --1.0 kb--  
 ggtatttaaa tctttgaaaa ttgtgagatca gctataagtc ctttctctag gaaaaacaca 060  
 gATTGATA CACTAAAAT TGGAAAGGCTA TTTCCTATGA GTCCGTAGAC TCCAAAATAA 120  
 AAAATTCTGC TCTAAATAAA AATGGTTAA CCTTTCTACT GTTTCTTTG TCTGATAATA 180  
 ACTTCCAAAA AAATACCTAG CTCAAGGGTT AATATTTCAT AAATAGTTAC TTTTTTTTT 240  
 EXON 17 CATTTCAGG AAGTACATCT CAGAATCTG ATTCTGGAAC AGATTTGTCT TTCCCATGGA 300  
 r gSerThrSer GlnAsnLeuA spSerGlyTh rAspLeuSer PheProTrpI (517)  
 TTCTGAATGT GCTTAATTAA AAAGCCTTG ATTTCACAA AGTGATCGAA AGTTTATCA 360  
 leLeuAsnVa lLeuAsnLeu LysAlaPheA spPheTyrLy sValIleGlu SerPheIleL (537)  
 AACCGAGAAGG CAACTTGACA AGAGAAATGA TAAAACATT AGAACGATGT GAACATCGAA 420  
 ysAlaGluG1 yAsnLeuThr ArgGluMetI leLysHisLe uGluArgCys GluHisArgI (557)  
 TCATGGAAATC CCTTGATGG CTCTCAGTAA GTAGCTAAAT AATTGAAGAA ATTCAATTCA 480  
 leMetGlySe rLeuAlaTrp LeuSer (565)  
 GTGCATATGG CAAACAAATT ATTGTTAGTG AGAGGTGTTT CTTAACAAAT CTACCTCAAG 540  
 AACAAATAGG GAATTAAATG AATAATGTTA TTTCAGTCTA TAGCCCCAGG ATCAagtggaa 600  
 atattagaat ggagcttaa tcgagcaccc taaccatct aatacagcnc agtgatttat 660  
 ttaagaatag cttttcttaa sacatgccac ttt 693  
 --70.0 kb--

12/18

## FIG. 6-6

EXON 18

```

attttcaat ataaagcggttgg aagggtatac atttttctac ttttttgtgt gtgggaagta 060
caaaaattgt CAATTGGGAA TTTCGAAGTA GAGAAAAATA TTTCATTCTG ACTTTAAAT 120
TGCCACTGTC AATTGTGCCCT AAAATTACATA GTACTTACCA TGTCAAACAA TATGATTTG 180
ATATGTACCT GGGAAAATTA TGCTTACTAA TGTTGTTTA ATTTCATCAT GTTICATATA 240
GGATTCACCT TTATTTGATC TTATTAACACA ATCAAAGGAC CGAGAAGGAC CAACTGATCA 300
AspSerPro LeuPheAspL euIleLysG1 nSerLysAsp ArgGluGlyP roThrAspHi (585)
CCTTGAATCT GCTTGTCCCTC TTAATCTTCC TCTCCAGAAT AACACACTG CAGCAGATAT 360
sLeuGluSer AlaCysProL euAsnLeuPr oLeuGlnAsn AsnHisThrA laAlaAspMe (605)
GTAAGCAAAA TATATGTTAT GTTGACCATC AAACIGCAAA TAGATTTAA GCATAAGTGC 420
AATGTAACAT TCTATTAAGA AAGTGTAGGG AATAGAATT TGAAATAAGAA TAGTTTCTGT 480
TTTAAGAAA TTAGTAATAA AAGGTACATg acccaaataa agtcataataa aagagtacag 540
agtgcactg aatcacctag gatttgcata atgagagcag ttttcatgg 589
--3.0 kb--

```

EXON 19

```

tgaaaaatccat ctggaaatcac cttatggctt caataccact ataattatta aaattgtACA 060
TTATACATAT ATAGCTTATTT TTTCATAATA AGGCAGTAAT CCCCAGGAAA AGCCATTAT 120
TAAAATAGAA TTAGATATGA TGATGACAAG CAGTTTCCT ATTAAATATAT CTTCAGC 180
TTGCATTTAA ATAGTCTGCT ATAATACCAA TTAAATAGAC AAGATGTATC TGGGTGTACA 240
ACCTTGAAGT GTATGTATAA TCTGTGATTG TTAGCCAATC TGAAATGAAG ACTTTCTT 300
TAAATATATC TAGGTATCTT TCTCCTGTAA GATCTCCAAA GAAAAAAGGT TCAACTACGC 360
tTyrLeu SerProValA rgSerProLy sLysLysGly SerThrThrA (621)
GTGTAAATTC TACTGCCAAAT GCAGAGACAC AAGCAACCTC AGCCTTCCAG ACCCAGAAC 420
rgValAsnSe rThrAlaAsn AlaGluThrG lnAlaThrSe rAlaPheGln ThrGlnLysP (641)
CATTGAAATC TACCTCTCTT TCACGTGTTT ATAAAAAAAGG TTAGTAGATG ATTATTTCA 480
roLeuLysSe rThrSerLeu SerLeuPhe yrLysLysV (654)
AGAGCATGGA CTCTGAAACT AGGCTGACTG GGTCATAATC ATGTTCTTC TACTTCTAG 540
GTACATTACT GGGCAAGTCA CTTAATATCT CTGTGTCTCA GTTCCTcat ctataaaatg 600
gaaatgataa tggtgcgaga tctttcttga ctattcagag tcgtttctg 650
--2.8 kb--

```

EXON 20

```

aaggaaaatc catccccatc cgggacatgc ctgnccctcg catttctca tctgttatccc 060
ttgtatatcg cctcataata aaccagtaaa catGTTTCTC TGGGGGAAAG AAAAGAGTGG 120
TAGAAAAGAG GTTTCTGTTA AAATGCTACT TAACAGCATT ATAATTAGTG TAATTCATG 180
ATTGAAAAAA AATCTACTTG TAATTCAAAA TGAACAGTAA AAATGACTAA TTTTCTTAT 240
TCCCACAGTG TATCGGCTAG CCTATCTCCG GCTAAATACA CTTGTGAAC GCCTCTGTC 300
al TyrArgLeuA laTyrLeuAr gLeuAsnThr LeuCysGluA rgLeuLeuSe (671)
TGAGCACCCA GAATTAGAAC ATATCATCTG GACCCTTTTC CAGCACACCC TGCAGAATGA 360
rGluHisPro GluLeuGluH isIleIleTr pThrLeuPhe GlnHisThrL euGlnAsnG1 (691)
GTATGAACTC ATGAGAGACA GGCATTTGGA CCAAGTAAGA AAATCAAGCA CTTCACCTTC 420
uTyrGluLeu MetArgAspA rgHisLeuAs pGln (702)
TCTCCTCCCT ACTTACTTGT TAACTGATTT CTTTCTTCTT TTCTTTCTT CTTTCTTCT 480
TTCTTTCTT CTTTCTTCTT TTCTTTCTT CTTTCTTCTT TTCTTTCTT TTCTTTCTT 540
TTTTTTGAG ATAGAGTCTC ACTCTGTAC CCAGGCTGGA GTGCAGTGGC GCAATCTCGG 600
CTCACTGCAA CCTCCGGCTC CCAGGTCAAG TGATTCTCCT GCCTCAGCCT ccnaggagct 660
aggatacagg cgtgtaccac cacaccttgt taattttgt tatTTAGTAG agacagg 717
--4.0 kb--

```

SUBSTITUTE SHEET

13/18

## FIG. 6-7

<u>SacI</u>	caagagccaa agttaggts atttacaaac caggtgatca gtcctggata attgaggcctt 060 8gtgatttgC ATTTGTTCT TAAACACAC TTTGGGTTAA ACACCTCATG TAGACTTCA 120 AACT <u>GAGCTC</u> AGTATGGAAA GAAATAACTC TGAGATTAA ACCTTTCTT TTTGAGGCTA 180 AAAGAAAGAA AATGGTATT TTTAAGAACAAACCATGTAA ATAAAATTCT GACTACTTT 240 EXON 21 ACATCAATT ATTACTAGA TTATGATGTG TTCCATGTAT GGCATATGCA AAGTGAAGAA 300 I LeMetMetCy sSerMetTyr GlyIleCysL ysValLysAs (716) TATAGACCTT AAATTCAAAA TCATTGTAAC AGCATAACAAG GATCTTCCTC ATGCTGTTCA 360 nIleAspLeu LysPheLysI leIleValTh rAlaTyrLys AspLeuProH isAlaValG1 (736) GGAGGTAGGT AATTTCAT AGTAAGTTT TTTGATAAAAT CCATATCCAT AACATAACAT 420 nGlu (737) AGGTAAATTCA TTTGATCTCA TTTATCATTA ATGAGATCAT ATATTCTGTC TGACCTTATT 480 ATGTAAATTCA ACAAAATAAAA ACTTTTATAT TATTATTTG TAACCTAAAT AGAATTGGAA 540 AGATAAGGGT AATTATGAAA TTACCCATAT CATACTTTT TATAAAGTTA ATAAATAATA 600 TTTATCCCT GTAATAAGCA GGTATTGTA ataaacttga catgagtcat agaacattag 660 atatcttgag 670
	--0.2 kb--
<u>XbaI</u>	tccatctgct gctgcctggc tatttctetc aatcgattct gtgacatttc acttctagaa 060 gagcaGCTAT AATCCAAGCC TAAGAAGTAA TTTTATTAT TTATTATTTT TTCCCTTATA 120 ATATGTGCTT CTACCAGTC AAAAGTATT ATAAACTATT AGAAAAGAAA ATCTAAAGGT 180 AGAAATTTA AAATTCAATT AACAAAGTAAA TTTTACTTTT TTTTTTTTTT TTTTTTTTTT 240 EXON 22 ACTGTTCTTC CTCAGACATT CAAACGTGTT TTGATCAAAG AAGAGGAGTA TGATTCTATT 300 ThrPh eLysArgVal LeuIleLysG luGluGluTy rAspSerIle (752) ATAGTATTCT ATAAGTCGGT CTTCATGCAG AGACTGAAAA CAAATATTTT GCAGTATGCT 360 IleValPheT yrAsnSerVa lPheMetGln ArgLeuLysT hrAsnIleLe uGlnTyrAla (772) TCCACCAGGG TAGGTCAAAA GTATCCTTIG ATTGGAAAAA TCTAATGTAA TGGGTCCACC 420 SerThrArg (775) EXON 23 AAAACATTAA ATAAATAATC TACTTTTIG TTTTGCTCT AGCCCCCTAC CTTGTCACCA 480 ProProTh rLeuSerPro (781) ATACCTCACA TTCCCTCGAAG CCCTTACAAG TTCCCTAGTT CACCCCTTACG GATTCCCTGGA 540 IleProHisI leProArgSe rProTyrLys PheProSerS erProLeuAr gIleProGly (801) GGGAACATCT ATATTCACC CCTGAAGAGT CCATATAAAA TTTCAGAAGG TCTGCCAACAA 600 GlyAsnIleT yrIleSerPr oLeuLysSer ProTyrLysI leSerGluG1 yLeuProThr (821) CCAACAAAAA TGACTCCAAG ATCAAGGTGT GTGTTTCTC TTTAGGGAAAG TAGTAAAGAA 660 ProThrLysM etThrProAr gSerAr (830) TGAGAGGGGG ATTATTTGA TCCAAGAATA AAAATATAA AGCATTCTTC ATTCAAATA 720 AGCTAGACTC TTGAAACTCT ATTGCTTAT TTAAGTAACA TAATAAGAAT ATGGGGGGGG 780 GGTGAAGAAA ATCTATTAC GACTTAAGCA ACGCAAGATG GCCGAATAGG AACAGCTCCG 840 gtctacagct cccagegtga gcacgcagaa gacgggtat ttctgcattt ccatctgagg 900 taccgggttc atctcaactag ggagtgccag acagt 935
	--7.4 kb--
	ttgataactt acccattgat ttatgaagaa ctaagttagg gtaacattga aacttgcctt 060 tgcctccct aaatatggc aatggcagna tatgttctg cagacctata actttgcTT 120 TAAAACAAAG AGACTAGGTG AGTATATGAT TAGACGGGCA CTGTTAGAAT AATTCCCAAA 180 TGAATATAAGT TTGTCAGTGG TTCTAGGGTA GAGGTAACTT TTAATTGGT ATTCTAATA 240 GTTCAGAATG ATGTATTAT GCTCATCTCT GCAAAATTGT ATATGGTTT TTATTACTAA 300 EXON 24 TTGGTATTTC ATCTTAACCT GACAGAATCT TAGTATCAAT TGGTGAATCA TTGGGGGTGA 360 gIleL euValSerIl eGlyGluSer PheGly (840)
<u>HindIII</u>	GTATTTCTT TCTATGAAAT ATAATAGTAT GCATTGTAAG TATAAAAGAA ATTAAAGCTT 420 TCTATAATT GAATTCCAA ATGCAGTTAT TCAAACACCT CATCCAGGCA TATTGCATAG 480 AATTATGA GATATATATA TCTCAGATT ACTTCAAAAT CAAGTTAAAT CTCAAATCAT 540 ACTCCTAATT GGTGAACCTC AAAACTTTC TAAATATCCA CTTGAGATTA TATAATACAT 600 ATATACATT GTGTATATAC ATACATATAT ACGTGAGCTG TTTTGCTCA CAACATTCT 660 ATCACCAAAT GTGTGAGATT TTTTCTCAC CCAAATCTAT TCTTcaactc tctggtgctt 720 ctacaattca attcaattct gacactaatt acccagag 758
	--2.8 kb--

14 / 18

## FIG. 6-8

EcoRI  
**EXON 25** gggatggaa ttaggttagtt attctgattt ttAGATTTT CATATCTTT ATTGGTCCA 060  
 ATGAAGCAGA AAATTAAAT GAAGTTATTA CCTTGCTG ATTTTGACA CACCTCAAAC 120  
 TATAACTTGA GGTGCTAAC TATGAAACAC TGGCATTAA TGATTAAAG TAAAGAATT 180  
**EXON 25** TGTAATTGT AGACTTCTGA GAAGTTCCAG AAAATAAAATC AGATGGTATG TAACAGCGAC 240  
 ThrSerG1 uLysPheGln LysIleAsnG 1nMetValCy sAsnSerAsp (856)  
 CGTGTGCTCA AAAGAAGTGC TGAAGGAAGC AACCCCTCTA AACCACTGAA AAAACTACGC 300  
 ArgValLeuL ysArgSerAl aGluGlySer AsnProProL ysProLeuLy sLysLeuArg (876)  
 TTTGATATTG AAGGATCAGA TGAAGCAGAT GGAAGGTAGG AACCAAGTTT GAATGTTTC 360  
 PheAspIleG 1uGlySerAs pGluAlaAsp GlySe (888)  
 CAGTAGCTGA GATGGTCATC TGGGAATCC AGAGTCTCAG CACTGCTCCT GGCTTATACC 420  
 AATTCTTTC ATGCCAAGTT TATTGGAAG TTGTGAGAAT GGCTCAAAT AATAGATATG 480  
 AGTGTAGTGC AAAGTTAAA ACATCTTACA AATTGCATAC CAACATTAG TGAAGATATC 540  
 TAATAAACCC TGATCTTTT TACAAAGCTA TTGATAAAAT TTTGTTATTG TTAACATTAA 600  
 ATTAAAAAT GTTACTCTT GAAAATATT AACCACTGTA TTTGTTGAGA ACCACTGAAA 660  
 AAATACATAG CATCATAAAAT TTGTGACATT TATGTTTAG ATGGTTAGTT TTTAAATT 720  
 AAAATTAAA GCTACTCACT AAAATAATAG CATAAAAGTAA GTCATCGAAA GCATCATAGT 780  
 TACTGGAAAT TTGAGTTTC CATTATAAA TACACATGAA ATGTTTGCA TTTTTTAAT 840  
**EXON 26** CTGCAGTAAA CATCTCCCAG GAGAGTCCAA ATTCAGCAG AAACCTGGCAG AAATGAGTAA 900  
 rLys HisLeuProG lyGluSerLy sPheGlnGln LysLeuAlaG 1uMetT (905)  
 GTACTTTTT CACCTTGTGT AAACGAAATA AACAAATTGTT TACACTGCAA GAAGTCTTT 960  
 CGTTATATAA AAGAATGTAT AATTCTTCA GTTGGAGGT TTGTTATGC ATTAAAATA 1020  
 TAATTCAATC AAGGTTATTG ATCTACAAAC ATTGTGGAT TAAATGTATG ATGAAAATG 1080  
 AAGGTCAATT TTACCCCTTC TATGATCTT CATGCAGGAA GACTAAGAAG TGAAACATTG 1140  
 CTTGACCACA TTCAAcacaa atggtacag tttagaaaata ctttagcaga actacaaaga 1200  
 ggaactattt gggagtgtta gatatagggaa aagttttat aaaccttagca tatgtaaaaac 1260  
 atcatcaccc ttattnaagg aataaccttt gattctaccg atttttaaac 1310  
 --1.7 kb--  
**EXON 27** tctagctatt tgaatatgca gtaaattaac tgtaactct acggtaactgt caaatactag 060  
 aatgaagacc acctttttt gcAAGGTCTT GAGGCCATC AGTTGACAT GAGCATAATA 120  
 TATATGGCAG CCACTTGCCA ACTAACCCAG TACCATGAAT GCTGTTAACAA GTTCTTCATC 180  
**EXON 27** CTTTTCCAG CTTCTACTCG AACACGAATG CAAAAGCAGA AAATGAATGA TAGCATGGAT 240  
 hrSerThrAr gThrArgMet GlnLysGlnL ysMetAsnAs pSerMetAsp (921)  
 ACCTCAAACA AGGAAGAGAA ATGAGGATCT CAGGACCTTG GTGGACACTG TGTACACCTG 300  
 ThrSerAsnL ysGluGluLy sEND 3 untranslated region begins here (928)  
 TGGATTCAATT GTCTCTACA GATGTGACTG TATAACTTTC CCAGGTTCTG TTTATGGCCA 360  
 CATTAAATAT CTTCAGCTCT TTTGTGGAT ATAAAATGTG CAGATGCAAT TGTTGGGTG 420  
 ATTCTTAAGC CACTTGAAAT GTTAGTCATT GTTATTATA CAAGATTGAA AATCTTGTGT 480  
 AAATCCTGCC ATTAAAAAG TTGTAGCAGA TTGTTTCCTC TTCCAAAGTA AAATGCTGT 540  
 GCTTTATGGA TAGTAAGAAT GCCCTAGAG TGGGAGTCCT GATAACCCAG GCCTGTCTGA 600  
 CTACTTTGCC TTCTTTGTA GCATATAGGT GATGTTGCT CTTGTTTTA TTAATTATA 660  
 TGTATATTGTTT TTTAATTAA CATGAACACC CTTAGAAAAT GTGTCCTATC TATCTTCCAA 720  
 ATGCAATTG ATTGACTGCC CATTCAACAA AATTATCCTG AACTCTCTG CAAAATGGA 780  
 TATTATTAGA AATTAGAAA AAATTACTAA TTTTACACAT TAGATTAT TTTACTATTG 840  
 GAATCTGATA TACTGTGTG TGTTTTATA AAATTTGCT TTTAATTAAA TAAAAGCTGG 900  
 AAGCAAAGTA TAACCATATG ATACTATCAT ACTACTGAAA CAGATTICAT ACCTCAGAAT 960  
 GTAAAAGAAC TTACTGATTA TTTCTTCAT CCAACTTATG TTTTAAATG AGGATTATTG 1020  
 ATAGTACTCT TGGTTTTAT ACCATTAGA TCACTGAATT TATAAAGTAC CCATCTAGTA 1080  
 CTTGAAAAAG TAAAGTGTTC TGCCAGATCT TAGGTATAGA GGACCTAAC ACAGTATATC 1140  
 CCAAGTGCAC TTCTAAATGT TTCTGGGTCC TGAAGAATT AGATACAAAT TAATTCTACT 1200  
 CCATAAACAG ACTGTTAATT ATAGGAGCCT TAATTTTTT TTCAATAGAGA TTTGTCTAAT 1260  
 TGCATCTCAA AATTATTCTG CCCTCCTAA TTTGGGAAGG TTTGTGTTT CTCTGGAATG 1320  
 GTACATGTCT TCCATGTATC TTTGAACG GCAATTGTCT ATTATCTTT TATTTTTTA 1380  
 AGTCAGTATG GCTAACACT GGCATGTTCA AAGCCACATT ATTCTAGTC CAAAATTACA 1440  
HindIII AGTAATCAAG GGTCAATTATG GGTAGGCAT TAATGTTCT ATCTGATTG GTGCAAAAGC 1500

15 / 18

## FIG. 6-9

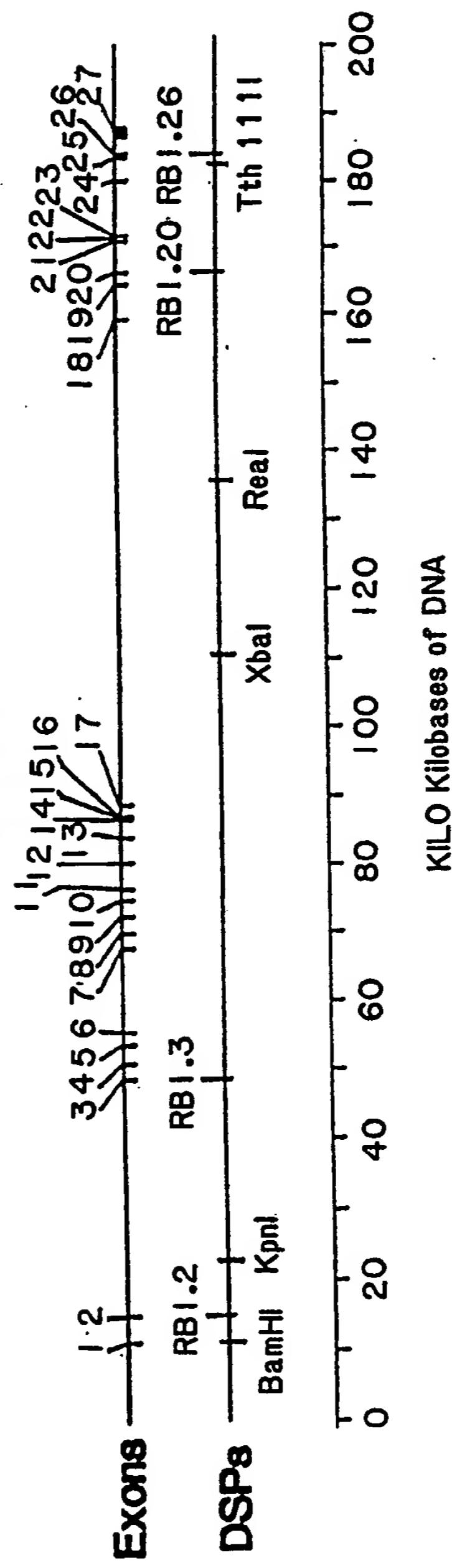
EXON 27 TTCAAATTAA AACAGCTGCA TTAGAAAG AGGCCCTCT CCCCTCCCT ACACCTAAGC 1560  
 (CONT'D) 3' untranslated region (continued)

GTGTATTAA ACTATCTGT GTGATTAAC TATTAGAGA TGCTGTAAC TAAAATAAGG 1620  
 GATATTAG GTAGCTTCAG CTAGCTTTA GGAAATCAC TTGCTCAAC TCAGAAATTAT 1680  
 TTTAAAG AATCTGGTC TTGTTAGAA ACAAAATTAT ATTGTGTGCT CATTAAAGT 1740  
 TCAAACCTAC TATTGAGA GTTATTGAG TAACAATGAC ACTAGAAAC TTGACTCCAT 1800  
 TTCACTCATG TTCTGCATG AATATCATAC AAATCAGTTA GTTTTAGGT CAAGGGCTTA 1860  
 CTATTTCTGG GTCTTTGCT ACTAAGTCA CATTAGAATT AGTGGCCAGAA TTITAGGAAC 1920  
 TTCAGAGTC GTGTATTGAG ATTCTTAA TAATGCTCA GATATTATG CTTTATTGCT 1980  
 TTITGTATT GGTTAAACT GTACATTAA AATGGCTATG TTACTATTCTACAAATTAA 2040  
 TAGTTGTCT ATTAAAT AATTTAGTG TAAAGAGTC TAAATGGTCTG ATGGTGTGTT 2100  
 polyadenylation signal sequence polyadenylation site  
 CTTTGATTA AGTACACTAA TGTCTCTTT TCTGTCTAGG AGAAGATAGA TAGAAGATAAA 2160  
 CTCTCCTAGT ATCTCATCCA TTCTAGCTT TAAGGGCT CTATATGCTA GAGATTCCA 2220  
 AATTATTTTC TTCAGCCCTG ATCTTTTCAC AGAGGTCAAG Gctttatag ccaacagaac 2280  
 tcttgcattc tactccccc taccccaatgt ctcccaatcc aactaaat caaaataaata 2340  
 aaaaatcttt tt 2352

SUBSTITUTE SHEET

16 / 18

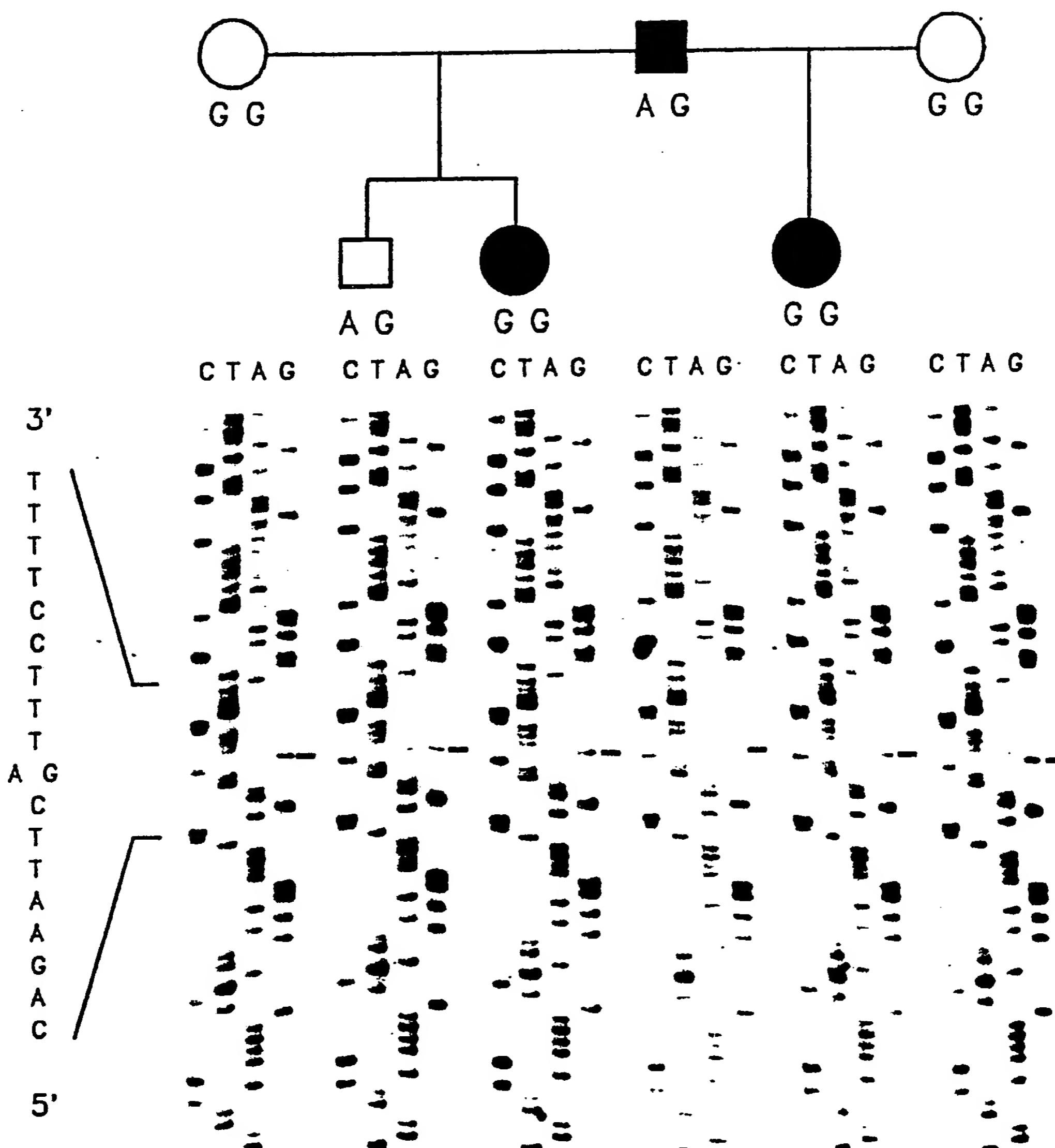
FIG. 7  
HUMAN RETINOBLASTOMA GENE



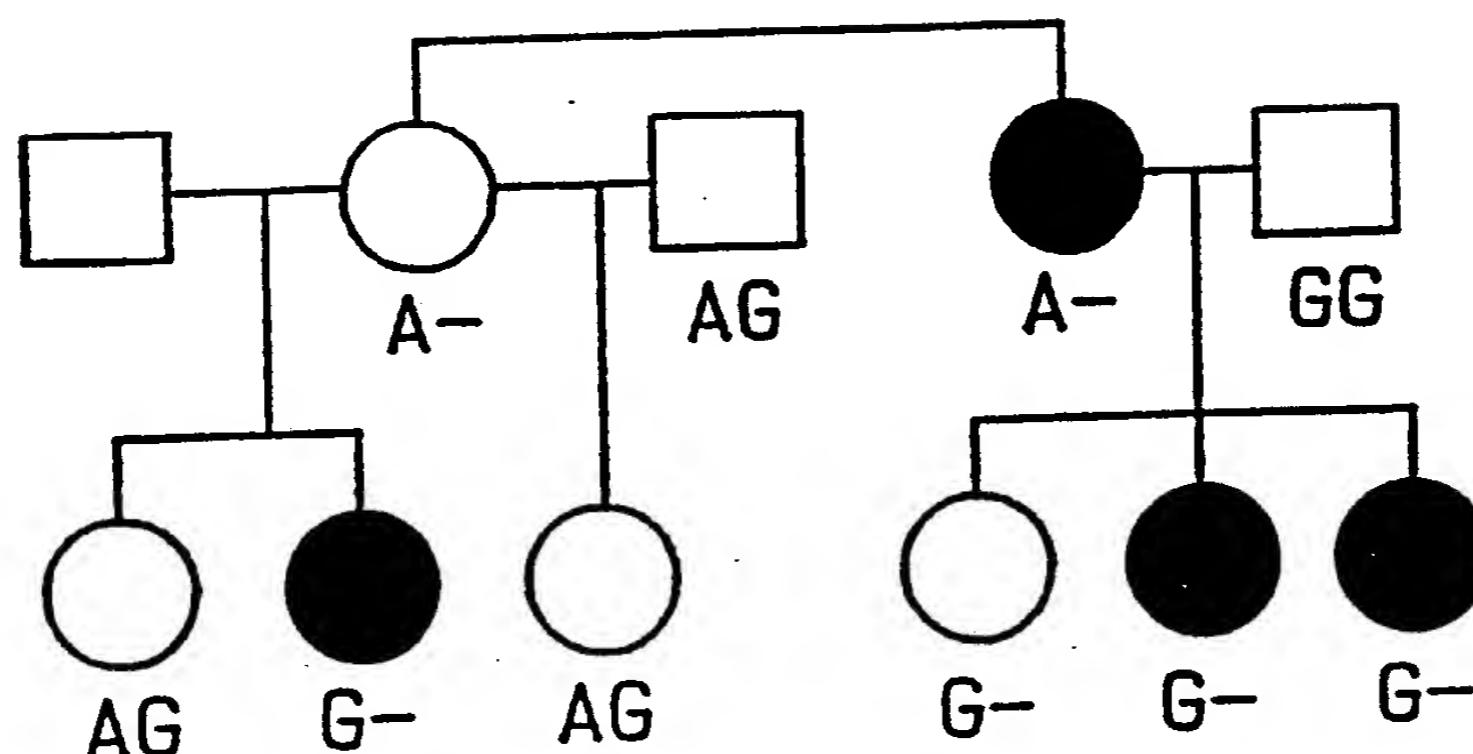
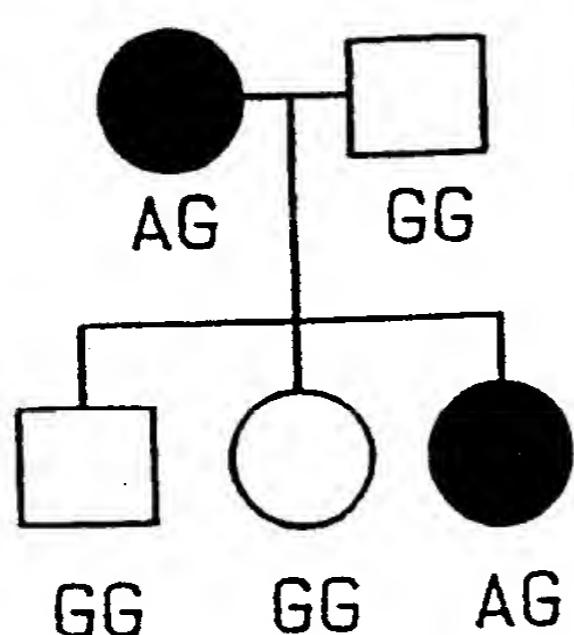
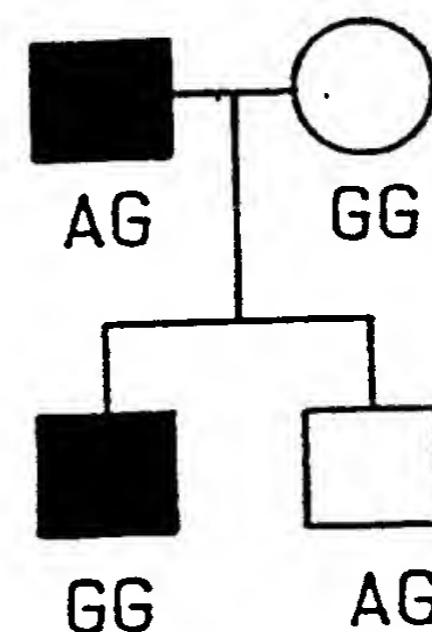
17/18

RB-6

FIG. 8



18/18

**FIG. 9 RB-32****RB-36****RB-50****SUBSTITUTE SHEET**

# INTERNATIONAL SEARCH REPORT

International Application No. PCT/US89/00293

## I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) <sup>6</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(4): C12Q 1/68 G01N 30/00  
U.S.: 435/6 536/27

## II. FIELDS SEARCHED

Minimum Documentation Searched <sup>7</sup>

Classification System	Classification Symbols
U.S.	435/6 536/27

Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in the Fields Searched <sup>8</sup>

APS; CAS; BIOSIS DATABASES

## III. DOCUMENTS CONSIDERED TO BE RELEVANT <sup>9</sup>

Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
X	SCIENCE, Volume 235 issued 13 March 1987 (Washington, D.C. USA) W. LEE "Human Retinoblastoma Susceptibility Gene: Cloning, Identification and Sequence" entire article.	1,4,8-16
P,X	EP, 0,259031 (Massachusetts Eye and Ear Infirmary) 09 March 1988	1,4,8-16
Y	Chemical Abstracts, Volume 107, No. 9 issued August 1987 (Columbus, Ohio, USA) S.A. LATT "DNA based detection of chromosome deletion & amplification" abstract No. 71579;, Cold Spring Harbor Symp. Quant. Biol. 51(1) 299-307	1,4,8-16
X Y	Chemical Abstracts, Volume 106 No. 7, issued February 1987 (Columbus, Ohio, USA) H. Scheffer, "A straightforward approach to isolate DNA sequences with potential linkage to the retinoblastoma locus" abstract No. 44946Z, Hum. Genet. 74(3) 249-255	<u>14</u> 8-13,15,16

\* Special categories of cited documents: <sup>10</sup>

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search

29 April 1989

Date of Mailing of this International Search Report

08 JUN 1989

International Searching Authority

ISA/US

Signature of Authorized Officer

SCOTT A. CHAMBERS

## FURTHER INFORMATION CONTINUED FROM THE SECOND SHEET

Y

Chemical Abstracts, Volume 105, No.17,  
 issued October 1986 (Columbus, Ohio, USA)  
 E.Y. LEE, "Molecular cloning of the human  
 esterase D gene a genetic marker of  
 retinoblastoma" abstract No. 147520q,  
 Proc. Natl. Acad. Sci. U.S.A. 83(17)  
 6337-41.

14

V.  OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE<sup>1</sup>

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1.  Claim numbers ..... because they relate to subject matter<sup>12</sup> not required to be searched by this Authority, namely:

2.  Claim numbers ..... because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out<sup>13</sup>, specifically:

3.  Claim numbers ..... because they are dependent claims not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI.  OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING<sup>2</sup>

This International Searching Authority found multiple inventions in this international application as follows:

- I. Claims 1, 4 and 8-16
- II. Claims 2 and 3
- III. Claims 5, 17 and 18
- IV. Claims 6, 7 and 19-21

1.  As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2.  As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3.  No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

1, 4 and 8-16

4.  As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

## Remark on Protest

- The additional search fees were accompanied by applicant's protest.
- No protest accompanied the payment of additional search fees.

## III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
Y	Chemical Abstracts, Volume 102, No. 7 issued February 1985 (Columbus, Ohio, U.S.A.) M. LALANDE "Isolation of human chromosome 13-specific DNA sequences cloned from flow sorted chromosome and potentially linked to the retinoblastoma locus" abstract No. 57143C, Cancer Genet. Cytogenet 13(4) 283-95.	1,4,8-16